Characterisation of the Bifunctional Aspartate Kinase: Diaminopimelate Decarboxylase from *Xylella fastidiosa*

Emma Kathryn Dorsey

A thesis submitted in partial fulfilment of the requirements

for the Degree of

Master of Science in Biochemistry

in the University of Canterbury

2014

School of Biological Sciences

University of Canterbury

Abstract

Xylella fastidiosa is a small, xylem-limited bacterium that causes a number of diseases in over 100 species of plants. Many of the species infected are economically important (such as coffee, grapevines, citrus, and almond) and billions of dollars worldwide are lost annually due to *X. fastidiosa* infection of crops. The bacterium colonises both plant and insect hosts, using the insect host to transfer it from plant to plant. Sequencing of the *X. fastidiosa* genome in 2000 discovered that while the genome is reduced, it contains a high number of putative bifunctional enzymes. One of these enzymes, aspartate kinase:diaminopimelate decarboxylase (AK:DapDc), occurs in only a handful of species and is predicted to catalyse the first and last steps of lysine biosynthesis. This study reports the first experimental characterisation of this enzyme. AK:DapDc was over-expressed in the pET30dSE plasmid in *Escherichia* coli BL21 DE3 cells. It was purified by Ni^{2+} His-Trap chromatography followed by size exclusion chromatography. Homology models of AK:DapDc were created in SWISS-MODEL, which indicate homology with the aspartate kinase from *Arabidopsis thaliana* and the diaminopimelate decarboxylase from *E. coli*. Circular dichroism, and analytical ultracentrifugation were used to obtain information about the secondary and quaternary structure of AK:DapDc. This data, in combination with the homology models, suggests that AK:DapDc exists as a dimer or tetramer in solution. A coupled enzyme assay to assay for diaminopimelate decarboxylase activity has been set up, and preliminary crystal screens have been carried out.

Acknowledgements

I would like to thank my supervisors Renwick Dobson and Sarah Kessans for their invaluable assistance, discussions and encouragement throughout the course of this study. The motivation and direction you provided were much appreciated. Thanks to Ren for asking the tough questions and encouraging me to further my understanding of my project. Thanks to Sarah for showing me the ropes in the lab, for the coffee breaks, motivation, support and inspiration.

I would like to thank the Perugini lab at LaTrobe University, Melbourne for sending me DNA. To Martin in particular, thanks for all the information and tips you gave me, and all the questions you answered.

Special thanks to my fellow sixth floor students for answering any and all questions, and for taking time out of your day to help me. Thanks to the Dobson lab group and everyone who helped me in the lab and proofread my thesis. In particular, thanks to Kat Donovan for your help with assays, Rachel North for helping me with secret purifications, and Moritz Lasse for help with CD and AUC, and testing buffers ‘for science’. Thanks especially to Letitia Gilmour for putting up with sharing a lab bench with me for 16 months.

Finally, this thesis would not have been possible without the continual encouragement and support of my family and friends. Thanks to my parents for their unconditional support and for listening to my rants about my project and science in general. Thank you so much dad for all your help with proofreading and formatting, and for being a source of calm in the storm. Thank you to mum for all your support, hugs and motivation.

Thanks to my flatmates for listening and for all your understanding during the last few months. To Letitia and Sian, thanks for the lunches and coffees, without which I would have lost my sanity long ago.

Table of Contents

Abbreviations

3D	three-dimensional
A	absorbance
ACT	Aspartate kinase: Chorismate mutase: TyrA
ADP	adenosine diphosphate
AEC	anion exchange chromatography
AK	aspartate kinase
Amp	ampicillin
ASADH	aspartic-semialdehyde dehydrogenase
ATP	adenosine triphosphate
AUC	analytical ultracentrifugation
bp	base pairs
C	centigrade
CD	circular dichroism
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
cm	centimeter
c(M)	continuous mass distribution
c(S)	continuous size distribution
Da	Dalton
DAP	diaminopimelate
DapDc	Diaminopimelate decarboxylase
DNA	deoxyribonucleic acid
dsDNA	double-stranded DNA
dNTP	deoxynucleotide triphosphate
DSF	differential scanning fluorimetry
DTT	dithiothreitol

E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic acid
g	gram
gDNA	genomic DNA
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPP	histidinol-phosphate phosphatase
HSDH I	homoserine dehydrogenase I
HT[V]	high tension voltage
IGPD	Imidazole-glycerolphosphate dehydratase
IPTG	Isopropyl-1-β-thiogalactopyranoside
Kan	kanamycin
kDa	kilo Dalton
*k*cat	enzyme catalytic turnover
K_M	Michaelis-Menten constant
L	litre
LB	Luria-Bertani broth
m	milli
M	molar
max	maximum
MES	2-(*N*-morpholino)ethanesulfonic acid
meso-DAP	*meso*-diaminopimelate
µ	micro
n	nano
NaCl	sodium chloride
NAD^+	nicotinamide adenine dinucleotide (oxidized)
NADH	nicotinamide adenine dinucleotide (reduced)
NADP+	nicotinamide adenine dinucleotide phosphate (oxidized)
NADPH	nicotinamide adenine dinucleotide phosphate (reduced)

nm	nanometers
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
PEP	phosphoenolpyruvate
pg	preparatory grade
PLP	pyridoxal-5'- phosphate
PRA-PH:PRA-CH	Phosphoribosyl-pyrophosphatase: Phosphoribosyl-AMP-cyclohydrolase
rpm	revolutions per minute
S. cerevisiae	Saccharomyces cerevisiae
ScSDH	Saccharopine dehydrogenase from S. cerevisiae
SDH	Saccharopine dehydrogenase
SDS	sodium dodecyl sulfate
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	size exclusion chromatography
TAE	Tris-acetate-EDTA
TIM	Triosephosphate isomerase
Tris	tris(hydroxymethyl)aminomethane
V	Voltage
X. fastidiosa	Xylella fastidiosa
*Xf*AK:DapDc	bifunctional aspartate kinase: diaminopimelate decarboxylase from *Xylella fastidiosa*

List of Figures

List of Tables

1. Introduction

1.1 *Xylella fastidiosa* (subspecies *fastidiosa*)

1.1.1 Introduction to the *Xylella fastidiosa* Bacterium

Xylella fastidiosa is a small, rod-shaped, xylem-limited bacterium that causes a number of diseases in over 100 species of plants. These diseases include Pierce's disease of grapevines (*Vitis vinifera*), phony peach disease in peach (*Prunus persica*), citrus variegated chlorosis in citrus (*Citrus* species), and several leaf scorch diseases in plants such as almond, coffee, plum, oleander and alfalfa[1]. Worldwide, billions of dollars are lost annually due to *X. fastidiosa* infection of economically important crops[2]. Symptoms of infection include leaf scorching (Figure 1.1A), smaller or dehydrated fruit (Figure 1.1B), leaf shedding, plant dieback, slowed growth and eventual death of the plant[1].

A. B.

Figure 1.1 - Effects of Pierce's disease in grape, A. Leaf scorch caused by water stress, B. Dehydrated, shrivelled grapes on an affected grapevine.

Pierce's disease was first observed in the 1880's[3] and *X. fastidiosa* was first shown to be the causal agent in 1978[4]. The bacterium was described in 1987 and was named for its slow, fastidious growth habit[5]. Since then, many strains of *X. fastidiosa* have been

identified, partially sequenced, and classified into subspecies[6-10]. A comparative database established in 2012 provides information on all sequenced strains of *X. fastidiosa* (http://www.xylella.lncc.br)[11].

1.1.2 Plant Hosts of *Xylella fastidiosa*

X. fastidiosa strains have a very broad host range and can infect over 100 species of plants. A full list of hosts can be found at http://www.cnr.berkeley.edu/xylella/control/hosts.htm[12]. Pierce's disease-causing strains, such as *X. fastidiosa* subsp. *fastidiosa,* have a broad host range, but primarily cause disease in grapevines[13]. Some species of wild plants can act as reservoirs for the bacterium. These species are not affected by the bacterium, or are affected to a lesser extent, and the bacterium can remain dormant in these plants for a period of time[14]. Many reservoir species do not show symptoms of infection, but facilitate re-infection of nearby grapes[12]. Riparian grasses, sedges and trees growing next to vineyards have been identified as major reservoirs[14]. Several species of New Zealand native plants are predicted to be susceptible to infection by *X. fastidiosa* (e.g. pohutukawa - *Metrosideros excelsa*), or have the potential to act as reservoirs[15]. Monterey pine (*Pinus radiata*), a major plantation species in New Zealand, is predicted to act as a reservoir[16].

1.1.3 Insect Hosts of *Xylella fastidiosa*

While *X. fastidiosa* is largely thought of as a plant pathogen, both plant and insect hosts need to be colonized for *X. fastidiosa* to persist in an area[17]. Plants are infected by transfer of *X. fastidiosa* from an insect host to the plant. The bacterium dwells as a biofilm in the mouthparts and foregut of the insect vector until it feeds from a plant[18]. Feeding transfers *X. fastidiosa* to the plant where it migrates to the xylem and lodges, forming a biofilm. The most common vectors are the glassy winged sharpshooter and blue-green sharpshooter

(Figure 1.2), but other sharpshooter leafhoppers and spittlebugs are also vectors[19]. New Zealand leafhoppers and spittlebugs are predicted to be potential insect hosts for *X. fastidiosa*[12].

A. B.

Figure 1.2 - The most common insect vectors for X. *fastidiosa*: A. The glassy-winged sharpshooter. B. The blue-green sharpshooter.

1.1.4 Mechanism of Pathogenesis

There is still debate over the mechanism of pathogenesis of *X. fastidiosa*[1]. Some symptoms such as wilting or leaf scorch suggest *X. fastidiosa* may be producing toxins. *X. fastidiosa* may also be causing a change in levels of plant-growth regulators[1]. Nitrogen metabolism is disturbed in diseased plants[20]. The mechanism with the most supporting evidence infers that *X. fastidiosa* causes blockage of the xylem, leading to water distress, leaf scorch, and eventual death of the plants[10].

1.1.5 Amino Acid Biosynthesis Genes in Reduced *X. fastidiosa* Genome

In 2000, the complete genome of the 9a5c clone of *X. fastidiosa* was reported[6]. Since then, seven more *X. fastidiosa* strains have been fully sequenced[8, 11] and ten strains have been partially sequenced (http://www.ncbi.nlm.nih.gov/genome/genomes/173). Simpson *et al.* (2000)[6] were the first to publically report the full genome of a bacterial plant pathogen, and *X. fastidiosa* was found to have a reduced genome[6]. This occurs when a bacterium has a

significantly smaller genome compared to its ancestor. Bacteria with reduced genomes use their host organism's processes to replace the lost functions[21]. This 'streamlining' of the bacterial genome may allow bacteria to replicate more often and use less energy, with the result being a more simple genome[22]. In contrast, bifunctional enzymes increase the complexity of genomes. Annotation of the 9a5c clone genome by Simpson et al. (2000) via BLAST and tRNAscan-SE[7] noted that, in addition to having a reduced genome, *X. fastidiosa* has nine predicted bifunctional enzymes, a high number compared to other gammaproteobacteria (Table 1.1). Four of these proteins were predicted to be involved in amino acid biosynthesis: aspartate kinase:diaminopimelate decarboxylase (AK:DapDc), aspartate kinase:homoserine dehydrogenase I (AK:HSDH I), phosphoribosyl-ATP-pyrophosphatase:phosphoribosyl-AMP-cyclohydrolase, and imidazole-glycerolphosphate dehydratase:histidinol-phosphate phosphatise. This is an interesting feature as many bacteria with reduced genomes have lost the genes necessary for biosynthesis of amino acids[21].

Table 1.1 - List of bifunctional enzymes predicted in *X. fastidiosa*'s genome.

Putative bifunctional enzyme	**Pathway**
aspartate kinase: diaminopimelate decarboxylase (AK:DapDc)	lysine biosynthesis
aspartate kinase: homoserine dehydrogenase I (AK:HSDH I)	threonine biosynthesis
phosphoribosyl-ATP-pyrophosphatase: phosphoribosyl-AMP-cyclohydrolase (PRA-PH-PRA-CH)	histidine biosynthesis
imidazole-glycerolphosphate dehydratase: histidinol-phosphate phosphatise (IGPD:HPP)	histidine biosynthesis
bifunctional purine biosynthesis protein (purH)	inosine 5-phosphate biosynthesis
bifunctional methylenetetrahydrofolate dehydrogenase/ methenyltetrahydrofolate cyclohydrolase (fold)	formyl THF biosynthesis
bifunctional penicillin-binding protein 1C precursor (pbpC)	peptidoglycan biosynthesis

Putative bifunctional enzyme	Pathway
dGTP-pyrophosphohydrolase/ thiamine phosphate synthase (mutt/thiE1)	DNA replication and repair/ thiamine (cofactor) synthesis
transcriptional repressor of the biotin operon/ biotin acetyl-CoA-carboxylase synthetase (birA/bioR/dhbB)	biotin metabolism

1.2 Bifunctional Enzymes

Bifunctional enzymes are one type of multifunctional enzyme. Enzymes can be multifunctional because they have promiscuous activity and recognise several substrates, for example cytochrome P450 mono-oxygenases which act on a wide range of toxic substrates[23]. Some multifunctional enzymes can produce multiple products from the same substrate, for example toluene 4-monooxygenase, which produces four different hydroxylation products from toluene[23]. Moonlighting enzymes are multifunctional because they catalyse a specialised reaction, but also have another role in the cell which is structural or regulatory, for example hexokinase from *Arabidopsis* species converts glucose to glucose-6-phosphate and also regulates a number of genes in response to intracellular glucose levels[24, 25].

Bifunctional enzymes are multifunctional because they have two distinct active sites on a single protein that catalyse different specialised reactions[26], which are often part of the same metabolic pathway. The two domains are encoded by one gene and are translated as one polypeptide. Bifunctional enzymes are found primarily in gammaproteobacteria and arise from the fusion of two genes encoding the two functions and the loss of any upstream promoter regions for one of those genes[27]. In theory, the presence of bifunctional enzymes can “allow increased coordination of gene expression”, but bifunctional enzymes are

actually quite rare despite this apparent advantage[24]. This is likely due to the fact that they arise as a result of random mutations which fuse two genes together.

Bifunctional enzymes can catalyse either two consecutive reactions, two non-consecutive reactions or two opposing reactions[26]. Catalysing two consecutive reactions can allow for substrate channelling from one active site to the other, reducing the loss of substrates to unwanted side reactions and diffusion. Many bifunctional enzymes that catalyse consecutive reactions do not exhibit channelling, suggesting this property is not as common as once believed[26]. Catalysing two non-consecutive reactions from the same pathway co-localizes the functions without having to form multi-protein complexes and can also be useful for regulation of the pathway[24]. The product of one reaction may bind allosterically to the enzyme and inhibit the other reaction quickly.

Another theory on the regulation and inhibition of bifunctional enzymes has recently been put forward. It is now being suggested that bifunctionality allows one domain/function to interact with the other function to cause a conformational change affecting activity and/or affinity, as opposed to the product of one domain's reaction allosterically inhibiting the other function[26]. Conformational changes that activate or deactivate domains may be transmitted from one domain to the other through an unstructured linker region which is in a particular conformational state. Interaction of one domain with the other, communicated *via* the linker, could be repressing the activity of the other site. Thus a linker between domains, and its conformational state, could be crucial for stabilizing domain(s). Some enzymes that have linkers, such as lysine-ketoglutarate reductase: saccharopine dehydrogenase of *Arabidopsis thaliana*, are sensitive to sodium chloride concentration[26]. The effect of sodium chloride on enzyme activity might be mediated by salt-induced changes in the linker. In this case, the linker may determine the nature of enzyme

conformational changes. One of the predicted bifunctional enzymes in *X. fastidiosa*, aspartate kinase:diaminopimelate decarboxylase, is predicted by the conserved domains database (CDD)[28] to have two separate conserved domains connected by a linker of approximately 30 amino acids. The state of the linker may affect the stability of one or both domains, making expression and purification more difficult, according to the above theory.

1.3 Lysine Biosynthesis

There are four known pathways that synthesise lysine from aspartate: the succinylase pathway, the acetylase pathway, the aminotransferase pathway, and the dehydrogenase pathway (Figure 1.3)[29]. The four pathways diverge after step four of the overall pathway. After the second reaction of lysine biosynthesis there is a branch point where the metabolite can be committed to lysine biosynthesis or funnelled into threonine, methionine or isoleucine biosynthesis. Meso-diaminopimelate can either be converted to lysine or used as a component of the bacterial cell wall. After the reaction carried out by dihydrodipicolinate synthase, the metabolite (*S*)-tetrahydrodipicolinate can be used by enzymes from any of the four pathways to synthesise lysine. These pathways also branch off into threonine, methionine, and isoleucine biosynthesis after the second step of the pathway, and to peptidoglycan biosynthesis before the last step. Plants utilise the succinylase or dehydrogenase pathways, whereas bacteria can use any of the four pathways[29]. Lysine biosynthesis genes (DapA-DapF) show a high level of co-expression, although they do not exist in an operon[27]. *X. fastidiosa* is predicted to synthesise lysine via the succinylase pathway[6].

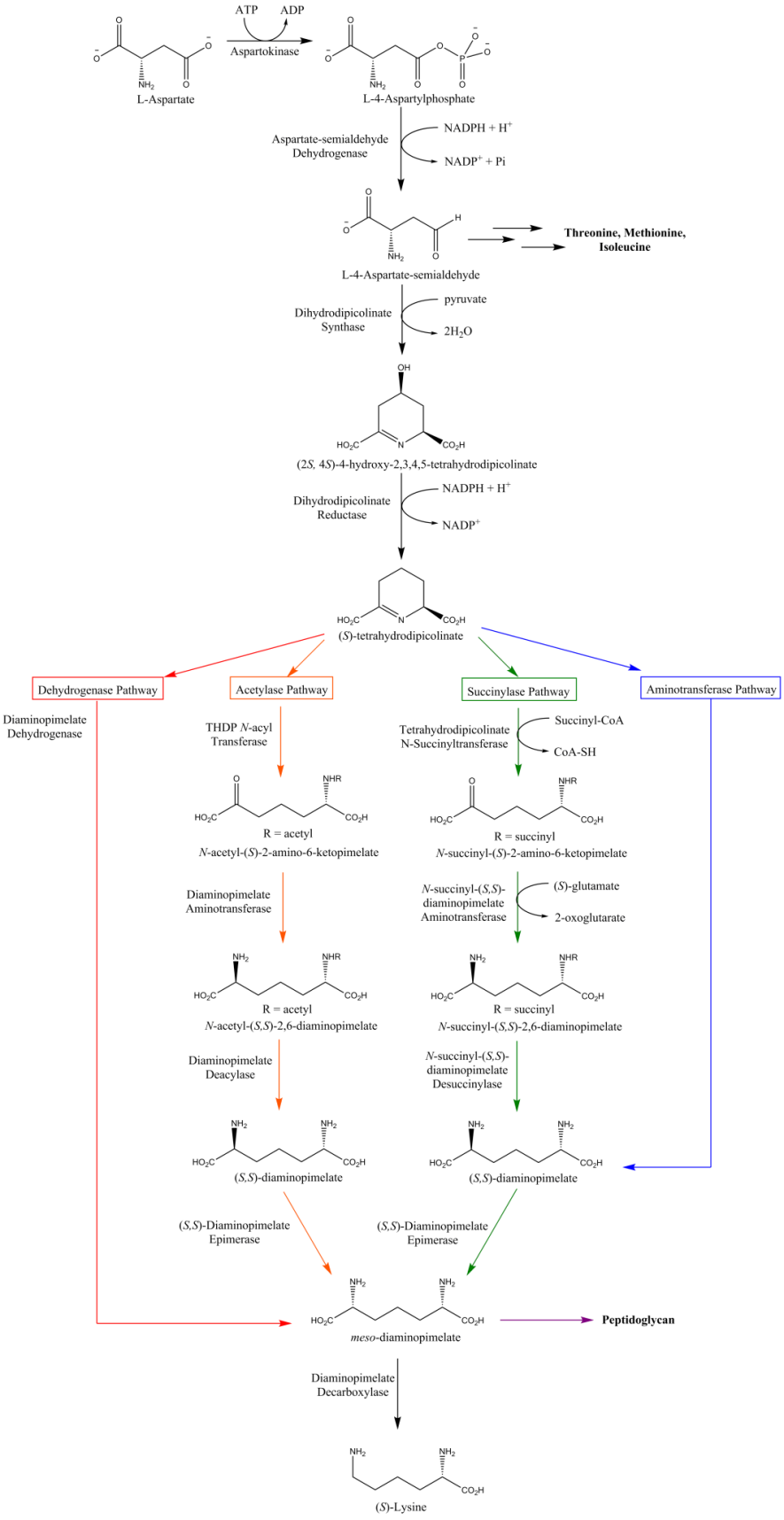


Figure 1.3 - Lysine biosynthesis pathways. Diagram drawn in Chemdraw.

Lysine biosynthesis is an essential pathway in *X. fastidiosa* subsp. *fastidiosa* and other xylem-dwelling bacteria. The xylem sap contains "a diversity of amino acids, organic acids, and inorganic ions"[30], but xylem-dwelling bacteria still need to synthesise a number of amino acids themselves. Xylem sap does not contain lysine, and *X. fastidiosa*'s insect hosts do not synthesise lysine either.

Essential lysine biosynthesis enzymes have the potential to be used as antibiotic targets, as they are not present in animals and thus have high specificity for bacteria and low toxicity in humans, animals, and plants[31, 32]. Due to the branched nature of lysine biosynthesis, inhibition of this pathway could affect all of these branches and prevent cell wall cross-linking. *Meso*-diaminopimelate is one of the D- amino acids that forms a short chain of amino acids to crosslink the rows of sugars (*N*-acetylglucosamine and *N*-acetylmuramic acid) in the cell wall. The reactions in lysine biosynthesis are not strictly essential to bacteria[32], but for the slow growing *X. fastidiosa* that may have evolved more specialized regulatory mechanisms, it could be essential and be a feasible antibacterial target.

Two of the predicted bifunctional enzymes in *X. fastidiosa*, AK:DapDc and AK:HSDH I, have roles in lysine biosynthesis. Both of these enzymes catalyse non-consecutive reactions (Figure 1.3), and occur at branch points in lysine biosynthesis, so their regulatory properties are of particular interest.

AK:HSDH I catalyses the first and third steps of lysine biosynthesis. The first step converts aspartate and adenosine triphosphate (ATP) to aspartyl 4-phosphate and adenosine diphosphate (ADP), committing aspartate to the aspartate-derived amino acid biosynthesis pathway. The third step of lysine biosynthesis is a branch point where the substrate can be directed into threonine synthesis or lysine synthesis. The HSDH I function commits the metabolite to threonine synthesis and is regulated by threonine. The AK:HSDH I gene is

more widespread/prevalent than AK:DapDc and occurs in many gammaproteobacteria[33, 34]. It is suggested that duplication of the AK gene (*ask*) was rapidly followed by fusion to the HSDH I gene (*hom)* as only *Vibrio* species of bacteria have two monofunctional *ask* genes (Figure 1.6)[27].

AK:DapDc is predicted to catalyse the first and last (ninth) steps of lysine biosynthesis. The AK domain catalyses the phosphorylation of aspartate to give aspartyl 4-phosphate. The last reaction converts meso-diaminopimelate to lysine and carbon dioxide, and occurs at a branch point between lysine biosynthesis and peptidoglycan synthesis. AK:DapDc could regulate the lysine synthesis pathway by feedback inhibition of the first step by the product of the last step. Its function is still putative and is only predicted to occur in a handful of species in the family Xanthomonadaceae. This study investigates the bifunctional AK:DapDc enzyme, as it is very rare and its function is still putative.

1.4 Aspartate kinase:diaminopimelate decarboxylase

The AK:DapDc gene from *X. fastidiosa* encodes the conserved domains for both functions, which are connected by a linker of approximately 30 amino acids[6]. The functions of this enzyme are still putative, that is, the functions are computationally predicted[35], but have not yet been experimentally confirmed. This gene is not found in an operon, unlike the other predicted bifunctional enzymes in *X. fastidiosa*[6, 27].

AK is an enzyme encoded by the *ask* or *LysC* gene. There are three types of aspartate kinases (I, II and III), with AK type III being the ancestral type[27]. Types I and II are bifunctional with HSDH I and II, respectively, in all gammaproteobacteria[27]. Each of these types is regulated by different end-product(s): type I is inhibited by threonine and isoleucine[36], type II is inhibited by methionine[33] and type III is inhibited by lysine[37]. The

aspartate kinase in AK:DapDc is a type III aspartate kinase[27] and is predicted to be allosterically inhibited by lysine.

AK catalyses the conversion of aspartate to L-4-aspartylphosphate (Figure 1.4). This reaction commits aspartate to aspartate-derived amino acid synthesis. The reaction is driven by the hydrolysis of ATP to ADP and requires a magnesium ion bound in the active site. The structures of 20 bacterial AKs have been fully or partially solved (PDB). Some research has focussed on just the regulatory domain of AK[38-40]. AK has an N-terminal catalytic domain and a C-terminal regulatory domain. The regulatory domain consists of multiple repeats of an aspartate kinase: chorismate mutase: TyrA (ACT) domain[41]. Each ACT domain has a binding site for an inhibitor and evolves independently of neighbouring ACT domains[42]. The catalytic domain has two lobes (N- and C-terminal lobes) and the active site of the enzyme lies between them[43].

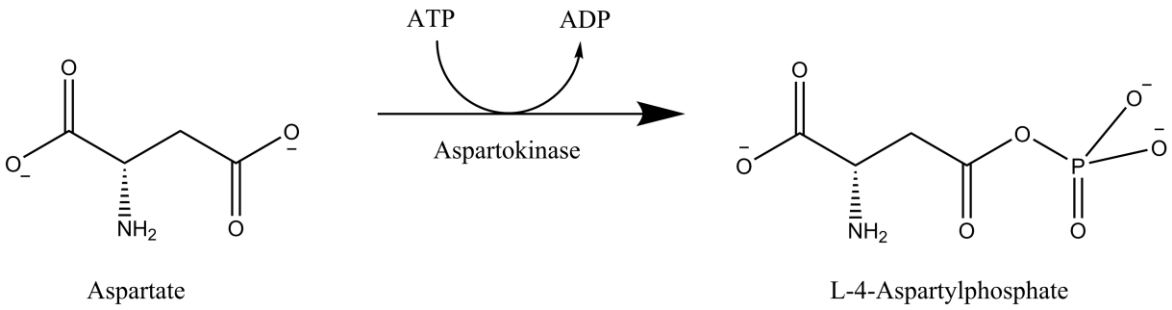


Figure 1.4 - Aspartokinase reaction: aspartate is phosphorylated by aspartate kinase, to produce L-4-aspartylphosphate.

Diaminopimelate decarboxylase (DapDc) is an enzyme encoded by the *lys*A gene. It takes *meso*-2,6-diaminopimelate and stereospecifically removes a carboxyl group to give *L*-lysine (Figure 1.5)[44]. DapDc requires pyridoxal 5-phosphate (PLP) as a cofactor. This is an unusual reaction for a PLP-dependent decarboxylase, as it stereospecifically converts a

D,L-enantiomer into an *L*-product, thus removing the *D*-stereocenter and resulting in inversion of configuration[45, 46]. DapDc is also allosterically inhibited by lysine.

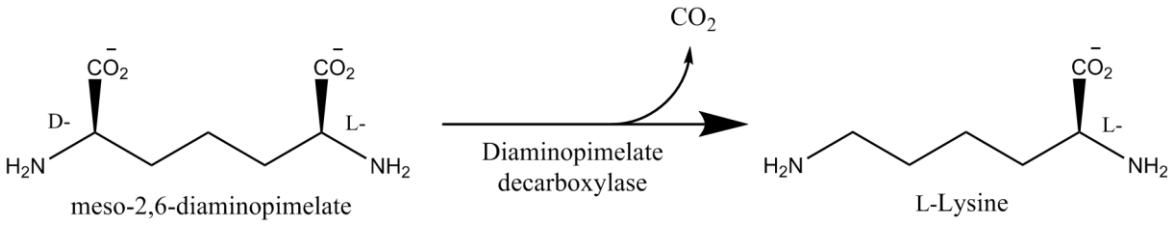


Figure 1.5 - Diaminopimelate decarboxylase reaction: the stereospecific decarboxylation of meso-2,6-diaminopimelate decarboxylase to produce L-lysine and carbon dioxide.

DapDc is a TIM-barrel enzyme with an additional β-sandwich domain. This makes it a fold type IV of PLP-dependent enzymes, which are homologs of alanine racemase, and are structurally similar to ornithine and arginine decarboxylases[47, 48]. DapDc is an obligate dimer, and dimerises in a "head-to-tail" fashion[49].The active site is located at the dimer interface and both subunits contribute residues to the active sites. Eleven bacterial and two archaeal DapDc enzyme structures have been crystallographically solved (PDB).

1.4.1.1 Proposed Mechanism of AK:DapDc Evolution

AK:DapDc is presumed to have evolved through random gene duplication of *ask* due to genetic drift (Figure 1.6)[27]. Only gammaproteobacteria contain multiple copies of the *ask* gene[27]. The *ask* gene duplicated in the gammaproteobacterial ancestor and one copy rapidly fused to the *hom* gene encoding HSDH to produce *ask-hom*. Species of gamma-proteobacteria group into two clusters: species in cluster one have only one *ask-hom* gene, while those in cluster two have two ask-hom genes due to a subsequent gene duplication[27]. The duplicate ask-hom gene was under less/no selection pressure compared to the original gene and this allowed for divergence of the two genes and specialisation[50]. In the

Xanthomonadaceae family, the monofunctional *ask* gene fused to *LysA*. *ask-LysA* contains the conserved domains encoding a type III AK. *X. fastidiosa's* other AK is a type I AK (fused to HSDH I).

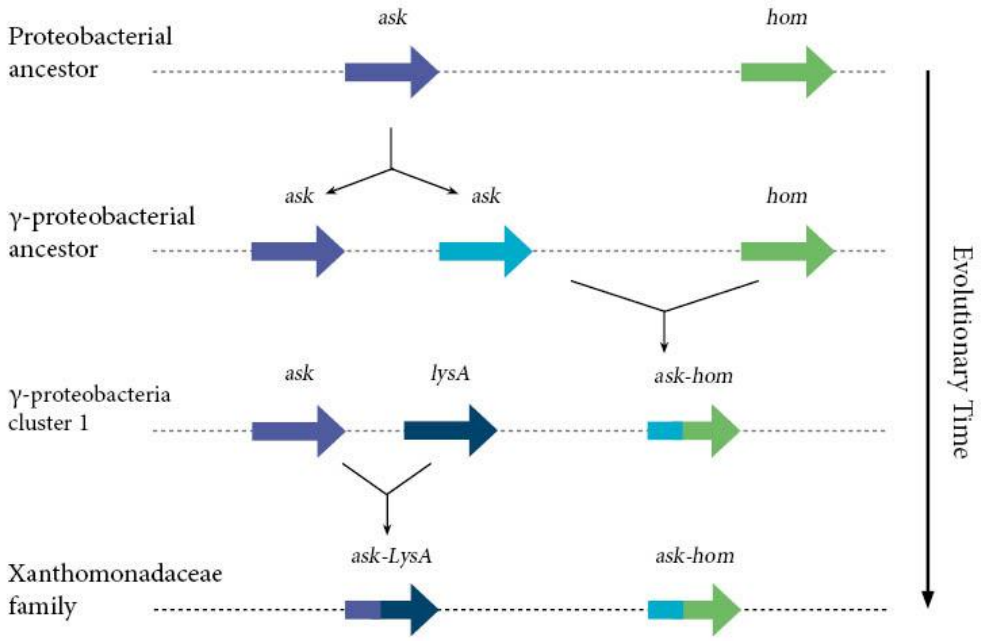


Figure 1.6 - Diagram of AK:DapDc (ask-LysA) evolution adapted from Fondi *et al.* (2007)[27]. The *ask* gene duplicated in the ancestral bacteria due to random mutation. One of the *ask* genes became fused to the *hom* gene encoding HSDH. In gammaproteobacteria (cluster 1) the fused *ask-hom* did not duplicate, but in the Xanthomonadaceae family the monofunctional *ask* gene fused to the *Lys*A gene encoding DapDc.

These gene duplication and gene fusion events allow for rapid diversification and exploration of sequence space[50, 51]. The ancestral, monofunctional AK may have been regulated to some degree by lysine, methionine, threonine, and isoleucine, but the regulation mechanism was unspecific[52]. Duplication, diversification and gene fusion has allowed the AKs to become sensitive to different signals, i.e. each is now regulated by a different end product, giving each an individual cellular role and fine-tuning the flux of aspartate-derived amino acid biosynthesis.

1.5 Scope of this Thesis

The overall aim of this thesis was to characterise the AK:DapDc enzyme from *X. fastidiosa* subsp. *fastidiosa*. This enzyme is of interest because it is one of many bifunctional enzymes encoded in the *X. fastidiosa* genome – an interesting feature in a reduced bacterial genome. Characterising this protein will add to the current knowledge base on enzyme bifunctionality and why this trait has evolved.

The first aim of this study was to over-express AK:DapDc and to purify the protein. The second aim was to biophysically characterise the protein using a range of methods that would investigate the secondary and quaternary structure of AK:DapDc, supported by computational data such as homology models. The third aim was to set up an activity assay in order to confirm the predicted functions of the enzyme.

I hypothesise that this enzyme may occur in solution as a dimer or tetramer, as both monofunctional forms of AK and DapDc are reported to occur as dimer and tetramer. I also hypothesise that AK:DapDc catalyses both predicted reactions, and is inhibited by lysine. If true, this would suggest that AK:DapDc regulates metabolite entry into lysine biosynthesis by inhibiting the first step in the pathway.

References

1. Hopkins, D. L. (1989) XYLELLA-FASTIDIOSA - XYLEM-LIMITED BACTERIAL PATHOGEN OF PLANTS, *Annual Review of Phytopathology 27*, 271-290.
2. Siebert, J. (2001) Economic impact of Pierce's disease on the California grape industry, In *California Department of Food and Agriculture Pierce's Disease Research Symposium*, pp 111-116.
3. Davis, M. J., Purcell, A. H., and Thomson, S. V. (1978) PIERCES DISEASE OF GRAPEVINES - ISOLATION OF CAUSAL BACTERIUM, *Science 199*, 75-77.
4. Davis, M. J., Purcell, A. H., and Thomson, S. V. (1978) Pierce's Disease of Grapevines: Isolation of the Causal Bacterium, *Science 199*, 75-77.
5. Wells, J. M., Raju, B. C., Hung, H.-Y., Weisburg, W. G., Mandelco-Paul, L., and Brenner, D. J. (1987) Xylella fastidiosa gen. nov., sp. nov: Gram-Negative, Xylem-Limited, Fastidious Plant Bacteria Related to Xanthomonas spp, *International Journal of Systematic Bacteriology 37*, 136-143.
6. Simpson, A. J. G., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M. C., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bove, J. M., Briones, M. R. S., Bueno, M. R. P., Camargo, A. A., Camargo, L. E. A., Carraro, D. M., Carrer, H., Colauto, N. B., Colombo, C., Costa, F. F., Costa, M. C. R., Costa-Neto, C. M., Coutinho, L. L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A. P., Ferreira, A. J. S., Ferreira, V. C. A., Ferro, J. A., Fraga, J. S., Franca, S. C., Franco, M. C., Frohme, M., Furlan, L. R., Garnier, M., Goldman, G. H., Goldman, M. H. S., Gomes, S. L., Gruber, A., Ho, P. L., Hoheisel, J. D., Junqueira, M. L., Kemper, E. L., Kitajima, J. P., Krieger, J. E., Kuramae, E. E., Laigret, F., Lambais, M. R., Leite, L. C. C., Lemos, E. G. M., Lemos, M. V. F., Lopes, S. A., Lopes, C. R., Machado, J. A., Machado, M. A., Madeira, A., Madeira, H. M. F., Marino, C. L., Marques, M. V., Martins, E. A. L., Martins, E. M. F., Matsukuma, A. Y., Menck, C. F. M., Miracca, E. C., Miyaki, C. Y., Monteiro-Vitorello, C. B., Moon, D. H., Nagai, M. A., Nascimento, A., Netto, L. E. S., Nhani, A., Nobrega, F. G., Nunes, L. R., Oliveira, M. A., de Oliveira, M. C., de Oliveira, R. C., Palmieri, D. A., Paris, A., Peixoto, B. R., Pereira, G. A. G., Pereira, H. A., Pesquero, J. B., Quaggio, R. B., Roberto, P. G., Rodrigues, V., Rosa, A. J. D., de Rosa, V. E., de Sa, R. G., Santelli, R. V., Sawasaki, H. E., da Silva, A. C. R., da Silva, A. M., da Silva, F. R., Silva, W. A., da Silveira, J. F., Silvestri, M. L. Z., Siqueira, W. J., de Souza, A. A., de Souza, A. P., Terenzi, M. F., Truffi, D., Tsai, S. M., Tsuhako, M. H., Vallada, H., Van Sluys, M. A., Verjovski-Almeida, S., Vettore, A. L., Zago, M. A., Zatz, M., Meidanis, J., Setubal, J. C., and Xylella fastidiosa Consortium, O. (2000) The genome sequence of the plant pathogen Xylella fastidiosa, *Nature 406*, 151-157.
7. Frohme, M., Camargo, A. A., Heber, S., Czink, C., Simpson, A. J. D., Hoheisel, J. D., and de Souza, A. P. (2000) Mapping analysis of the Xylella fastidiosa genome, *Nucleic Acids Research 28*, 3100-3104.
8. Bhattacharyya, A., Stilwagen, S., Reznik, G., Feil, H., Feil, W. S., Anderson, I., Bernal, A., D'Souza, M., Ivanova, N., Kapatral, V., Larsen, N., Los, T., Lykidis, A., Selkov, E., Walunas, T. L., Purcell, A., Edwards, R. A., Hawkins, T., Haselkorn, R., Overbeek, R., Kyrpides, N. C., and Predki, P. F. (2002) Draft sequencing and comparative genomics of Xylella fastidiosa strains reveal novel biological insights, *Genome Research 12*, 1556-1563.

9. Schaad, N. W., Postnikova, E., Lacy, G., Fatmi, M., and Chang, C. J. (2004) Xylella fastidiosa subspecies: X. fastidiosa subsp piercei, subsp. nov., X. fastidiosa subsp. multiplex subsp. nov., and X. fastidiosa subsp. pauca subsp. nov. (vol 27, pg 290, 2004), *Systematic and Applied Microbiology 27*, 763-763.
10. Schreiber, H. L., Koirala, M., Lara, A., Ojeda, M., Dowd, S. E., Bextine, B., and Morano, L. (2010) Unraveling the First Xylella fastidiosa subsp fastidiosa Genome from Texas, *Southwestern Entomologist 35*, 479-483.
11. Varani, A. M., Monteiro-Vitorello, C. B., de Almeida, L. G. P., Souza, R. C., Cunha, O. L., Lima, W. C., Civerolo, E., Van Sluys, M. A., and Vasconcelos, A. T. R. (2012) Xylella fastidiosa comparative genomic database is an information resource to explore the annotation, genomic features, and biology of different strains, *Genetics and Molecular Biology 35*, 149-152.
12. Janse, J. D., and Obradovic, A. (2010) Xylella fastidiosa: Its biology, diagnosis, control and risks, *Journal of Plant Pathology 92*, S35-S48.
13. Hopkins, D. L., and Purcell, A. H. (2002) Xylella fastidiosa: Cause of Pierce's disease of grapevine and other emergent diseases, *Plant Disease 86*, 1056-1066.
14. Costa, H. S., Raetz, E., Pinckard, T. R., Gispert, C., Hernandez-Martinez, R., Dumenyo, C. K., and Cooksey, D. A. (2004) Plant hosts of Xylella fastidiosa in and near southern California vineyards, *Plant Disease 88*, 1255-1261.
15. McGaha, L. A., Jackson, B., Bextine, B., McCullough, D., and Morano, L. (2007) Potential Plant Reservoirs for Xylella fastidiosa in South Texas, *American Journal of Enology and Viticulture 58*, 398-401.
16. Nunney, L., Vickerman, D. B., Bromley, R. E., Russell, S. A., Hartman, J. R., Morano, L. D., and Stouthamer, R. (2013) Recent Evolutionary Radiation and Host Plant Specialization in the Xylella fastidiosa Subspecies Native to the United States, *Applied and Environmental Microbiology 79*, 2189-2200.
17. Rodrigo Piacentini Paes De, A., and Adam Christopher, R. Xylella fastidiosa Diversity, pp 107-116, 台中市：農業試驗所.
18. Hill, B. L., and Purcell, A. H. (1995) ACQUISITION AND RETENTION OF XYLELLA-FASTIDIOSA BY AN EFFICIENT VECTOR, GRAPHOCEPHALA-ATROPUNCTATA, *Phytopathology 85*, 209-212.
19. Purcell, A. H. (1975) ROLE OF BLUE-GREEN SHARPSHOOTER, HORDNIA-CIRCELLATA IN EPIDEMIOLOGY OF PIERCES DISEASE OF GRAPEVINES, *Environmental Entomology 4*, 745-752.
20. Purcino, R. P., Medina, C. L., Martins, D., Winck, F. V., Machado, E. C., Novello, J. C., Machado, M. A., and Mazzafera, P. (2007) Xylella fastidiosa disturbs nitrogen metabolism and causes a stress response in sweet orange Citrus sinensis cv. Pera, *Journal of Experimental Botany 58*, 2733-2744.
21. Moran, N. A. (2002) Microbial Minimalism: Genome Reduction in Bacterial Pathogens, *Cell 108*, 583-586.
22. Mira, A., Ochman, H., and Moran, N. A. (2001) Deletional bias and the evolution of bacterial genomes, *Trends in Genetics 17*, 589-596.
23. Copley, S. D. (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity, *Current Opinion in Chemical Biology 7*, 265-272.
24. Moore, B. D. (2004) Bifunctional and moonlighting enzymes: lighting the way to regulatory control, *Trends in Plant Science 9*, 221-228.
25. Rolland, F., Moore, B., and Sheen, J. (2002) Sugar sensing and signaling in plants, *Plant Cell 14*, S185-S205.

26. Nagradova, N. (2003) Interdomain communications in bifunctional enzymes: How are different activities coordinated?, *Iubmb Life 55*, 459-466.
27. Fondi, M., Brilli, M., and Fani, R. (2007) On the origin and evolution of biosynthetic pathways: integrating microarray data with structure and organization of the common pathway genes, *Bmc Bioinformatics 8*.
28. Marchler-Bauer, A., Zheng, C., Chitsaz, F., Derbyshire, M. K., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Lu, S., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Zhang, D., and Bryant, S. H. (2013) CDD: conserved domains and protein three-dimensional structure, *Nucleic Acids Research 41*, D348-D352.
29. Velasco, A. M., Leguina, J. I., and Lazcano, A. (2002) Molecular evolution of the lysine biosynthetic pathways, *Journal of Molecular Evolution 55*, 445-459.
30. Purcell, A. H., and Hopkins, D. L. (1996) Fastidious xylem-limited bacterial plant pathogens, *Annual Review of Phytopathology 34*, 131-151.
31. Hutton, C. A., Perugini, M. A., and Gerrard, J. A. (2007) Inhibition of lysine biosynthesis: an evolving antibiotic strategy, *Molecular BioSystems 3*, 458-465.
32. Dogovski, C., Atkinson, S. C., Dommaraju, S. R., Hor, L., Dobson, R., Hutton, C., Gerrard, J. A., and Perugini, M. (2009) Lysine biosynthesis in bacteria: an unchartered pathway for novel antibiotic design, *Encyclopedia of life support systems 11*, 116-136.
33. Falcoz-Kelly, F., van Rapenbusch, R., and Cohen, G. N. (1969) The methionine-repressible homoserine dehydrogenase and aspartokinase activities of Escherichia coli K 12. Preparation of the homogeneous protein catalyzing the two activities. Molecular weight of the native enzyme and of its subunits, *European Journal of Biochemistry 8*, 146-152.
34. Falcoz-Kelly, F., Janin, J., Saari, J. C., Véron, M., Truffa-Bachi, P., and Cohen, G. N. (1972) Revised structure of aspartokinase I-homoserine dehydrogenase I of Escherichia coli K12. Evidence for four identical subunits, *European Journal of Biochemistry 28*, 507-519.
35. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *Journal of Molecular Biology 215*, 403-410.
36. Truffa-Bachi, P., van Rapenbusch, R., Janin, J., Gros, C., and Cohen, G. N. (1968) The Threonine-Sensitive Homoserine Dehydrogenase and Aspartokinase Activities of Escherichia coli K 12, *European Journal of Biochemistry 5*, 73-80.
37. Funkhous.Jd, Abraham, A., Smith, V. A., and Smith, W. G. (1974) Kinetic and molecular-properties of Lysine-sensitive Aspartokinase - Factors influencing Lysine-mediated association reaction and their relationship to cooperativity of lysine inhibition, *Journal of Biological Chemistry 249*, 5478-5484.
38. Yang, Q. Z., Yu, K., Yan, L. M., Li, Y. Y., Chen, C., and Li, X. M. (2011) Structural view of the regulatory subunit of aspartate kinase from Mycobacterium tuberculosis, *Protein & Cell 2*, 745-754.
39. Yoshida, A., Tomita, T., Kuzuyama, T., and Nishiyama, M. (2007) Purification, crystallization and preliminary X-ray analysis of the regulatory subunit of aspartate kinase from Thermus thermophilus, *Acta Crystallographica Section F-Structural Biology and Crystallization Communications 63*, 96-98.
40. Yoshida, A., Tomita, T., Kurihara, T., Fushinobu, S., Kuzuyama, T., and Nishiyama, M. (2007) Structural Insight into Concerted Inhibition of α2β2-Type Aspartate Kinase from Corynebacterium glutamicum, *Journal of Molecular Biology 368*, 521-536.

41. Robin, A. Y., Cobessi, D., Curien, G., Robert-Genthon, M., Ferrer, J. L., and Dumas, R. (2010) A New Mode of Dimerization of Allosteric Enzymes with ACT Domains Revealed by the Crystal Structure of the Aspartate Kinase from Cyanobacteria, *Journal of Molecular Biology 399*, 283-293.
42. Chipman, D. M., and Shaanan, B. (2001) The ACT domain family, *Current Opinion in Structural Biology 11*, 694-700.
43. Kotaka, M., Ren, J., Lockyer, M., Hawkins, A. R., and Stammers, D. K. (2006) Structures of R- and T-state Escherichia coli aspartokinase III - Mechanisms of the allosteric transition and inhibition by lysine, *Journal of Biological Chemistry 281*, 31544-31552.
44. Kelland, J. G., Palcic, M. M., Pickard, M. A., and Vederas, J. C. (1985) Stereochemistry of Lysine formation by meso-Diaminopimelate decarboxylase from wheat-germ - Use of H-1-C-13 NMR Shift correlation to detect stereospecific deuterium labelling, *Biochemistry 24*, 3263-3267.
45. White, P. J., and Kelly, B. (1965) Purification and properties of Diaminopimelate decarboxylase from Escherichia coli, *Biochemical Journal 96*, 75-&.
46. Fogle, E. J., and Toney, M. D. (2011) Analysis of catalytic determinants of diaminopimelate and ornithine decarboxylases using alternate substrates, *Biochimica Et Biophysica Acta-Proteins and Proteomics 1814*, 1113-1119.
47. Lee, J., Michael, A. J., Martynowski, D., Goldsmith, E. J., and Phillips, M. A. (2007) Phylogenetic diversity and the structural basis of substrate specificity in the beta/alpha-barrel fold basic amino acid decarboxylases, *Journal of Biological Chemistry 282*, 27115-27125.
48. Sandmeier, E., Hale, T. I., and Christen, P. (1994) Multiple evolutionary origin of pyridoxal-5'-phosphate-dependent amino acid decarboxylases, *European Journal of Biochemistry 221*, 997-1002.
49. Hu, T. C., Wu, D. L., Chen, J., Ding, J. P., Jiang, H. L., and Shen, X. (2008) The catalytic intermediate stabilized by a "Down" active site loop for diaminopimelate decarboxylase from Helicobacter pylori - Enzymatic characterization with crystal structure analysis, *Journal of Biological Chemistry 283*, 21284-21293.
50. Lynch, M. (2007) The frailty of adaptive hypotheses for the origins of organismal complexity, *Proceedings of the National Academy of Sciences 104*, 8597-8604.
51. Lynch, M., Bobay, L. M., Catania, F., Gout, J. F., and Rho, M. (2011) The Repatterning of Eukaryotic Genomes by Random Genetic Drift, In *Annual Review of Genomics and Human Genetics, Vol 12* (Chakravarti, A., and Green, E., Eds.), pp 347-366.
52. Jensen, R. A. (1976) Enzyme recruitment in evolution of new function, *Annual Review of Microbiology 30*, 409-425.

2. Methods and Materials

2.1 Experimental Reagents

2.1.1 Chemical Reagents

All buffers, reactions, and solutions requiring water utilised pyrogen-free, DNase-free, ultrapure type 1 Milli-Q water dispensed from a Milli-Q Reference fitted with a BioPak Polisher (Merck Millipore, Billerica, Massachusetts, USA). Unless specifically stated otherwise, general chemicals used to make buffers and solutions were purchased from a number of suppliers including Amersham Biosciences, Bioline, Life Technologies, Merck, Sigma-Aldrich, and Thermo Fisher Scientific.

2.1.2 Biological Materials

Xylella fastidiosa subsp. *fastidiosa* (ICMP 15197) genomic DNA (25 ng/μL) was sourced from Landcare Research. Primers were supplied by Integrated DNA Technologies (IDT, Singapore). The His-tagged *Saccharomyces cerevisiae* saccharopine dehydrogenase (SDH) gene in the pET16b vector was supplied by Matthew Perugini's lab group at LaTrobe University (Melbourne, Australia). Restriction enzymes and their associated buffers were supplied by New England BioLabs Inc. (NEB, Ipswich, Massachusetts, USA). The vector pCR™2.1-TOPO®, part of the pCR2.1TOPO® TA Cloning® Kit used for cloning, was supplied by Invitrogen (Invitrogen™ by Life Technologies™, Carlsbad, California, USA). Lysozyme from chicken egg white was obtained from Sigma-Aldrich (St. Louis, Missouri, USA).

2.1.3 General Equipment

All eppendorf tubes and pipette tips used throughout all experiments were sterilised by autoclaving at 121°C for 20 minutes and were allowed to cool prior to use. Large cultures were centrifuged in the Sorvall™ RC 6 Plus Centrifuge (Thermo Scientific, Waltham, Massachusetts, USA) using the fixed angle F10S-6x500y rotor, while samples of 50 mL or less were centrifuged using the Eppendorf Centrifuge 5810R (Eppendorf, Hamburg, Germany) with the fixed angle rotors F34-6-38 and F45-30-11.

A Minitron incubator shaker (Infors HT, Bottmingen, Switzerland) was used for the incubation of all bacterial cell cultures. DNA concentrations were measured using a Nanodrop Spectrophotometer ND-1000 (Thermo Scientific, Wilmington, Delaware, USA). Buffers were filtered either through a Whatman™ Klari-Flex™ 250 mL 0.22 μm PES Filter Funnel (Fisher Scientific, Loughborough, UK) or through a 33 mm Millex-GV 0.22 μm PVDF syringe filter (Merck Millipore, Billerica, Massachusetts, USA). Finally, all pH measurements were performed using the UltraBasic pH/mV meter (Denver Instrument, Bohemia, New York, USA).

2.1.4 Standard Solutions

Antibiotics were made up as concentrated stock solutions. Kanamycin (kan) (kanamycin sulphate, Roche, Indianapolis, Indiana, USA) stock solutions were made up in water at 100 mg/mL, filter sterilised with a 0.22 μm filter and used in media at a final concentration of 100 μg/mL. Ampicillin (amp) (ampicillin sodium salt; Applichem, Darmstadt, Germany) was prepared in water at 100 mg/mL, filter sterilised with a 0.22 μm filter and used in media at a final concentration of 100 μg/mL. Chloramphenicol (cam) (Roche, Indianapolis,

Indiana, USA) was prepared in 100% ethanol at 30 mg/mL, filter sterilised with a 0.22 μm filter and used in media at a final concentration of 30 μg/mL.

Diaminopimelate (Sigma-aldrich, D1377- a mixture of LL-, DD-, *meso*-) was made as a concentrated stock solution of 20 mM in water. L-lysine was made as a stock of 40 mM in water. Pyridoxal phosphate stocks were made up in water to a concentration of 20 mM. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was made up in water as a stock of 1 M.

All stock solutions were stored at -20°C, unless otherwise specified.

2.2 Molecular Biology

2.2.1 Bacterial Strains, Plasmids, Genes, and Primers

2.2.1.1 Bacterial Strains

Escherichia coli XL-1 Blue (standard laboratory strain) cells were used for cloning. *E. coli* BL21 (DE3) cells were used throughout this study for protein expression. *E. coli* BL21 (DE3) Rosetta cells were used in expression trials. The Rosetta cells contained the pRARE plasmid, which confers chloramphenicol resistance and contains rare tRNA genes.

2.2.1.2 Plasmids

Cells were transformed with vectors listed in Table 2.1, containing either the gene for AK:DapDc (Table 2.3) or SDH. Plasmids pCR2.1-TOPO-XF, pBluescript-XF, pET30ΔSE-HisXF, and pET30ΔSE-XF were constructed using the protocols described in sections 2.2.4-2.3.1 below. Maps for these plasmids are found in Appendix II.

Table 2.1 - List of plasmids used in this study.

Plasmid Name (Parent plasmid)	**Purpose**	**Antibiotic Resistance**	**Insert**
pCR2.1-TOPO-XF (pCR2.1-TOPO)	Cloning vector	kan^R, amp^R	XFLM_07435 - *X. fastidiosa* AK:DapDc gene
pBluescript-XF (pBluescript SK^+)	Cloning vector	amp^R	XFLM_07435 - *X. fastidiosa* AK:DapDc gene
pET30ΔSE-XF (pET30ΔSE)	Expression vector	kan^R	XFLM_07435 - *X. fastidiosa* AK:DapDc gene
pET30ΔSE-HisXF (pET30ΔSE)	Expression vector	kan^R	XFLM_07435 – His-tagged *X. fastidiosa* AK:DapDc gene
pET16b-SDH (pET16b)	Expression vector	amp^R, cam^R	*Saccharomyces cerevisiae* Saccharopine Dehydrogenase (SDH) gene

2.2.1.3 XFLM_07435

The AK:DapDc gene (XFLM_07435) sequence was determined by sequencing (Table 2.3) and is identical to other AK:DapDc genes from *X. fastidiosa* subsp. *fastidiosa* strains already sequenced (e.g. *X. f. fastidiosa* GB514).

2.2.1.4 Primers

The primers used in this study (Table 2.2) were obtained from Integrated DNA Technologies (IDT, Singapore). All primers were resuspended in MilliQ water to a final concentration of 100 μM. Primer pair oRD53(F) and oRD54(R) were used to clone a pET30ΔSE-HisXF construct with an N-terminal His-tag. The primer pair oRD55(F) and oRD54(R) were used to engineer *Nde*I and *Hind*III restriction sites to create a pETΔSE-XF

construct with no tags. The forward primer oRD56(F) anneals approximately one third of the way along the XFLM_07435 and was used in conjunction with the T7 vector primers to sequence the XFLM_07435 gene. Two primers are not sufficient to sequence the entire gene accurately.

Table 2.2 - Primers used in this study.

Primer Name and Sequence	**Restriction Site**	**Melting Temperature**
oRD53(F): 5' GAG CTC ATG TCC GCT TCC CCT ACA GC 3'	*Sac*I	Tm= 64.6°C
oRD54(R): 5' AAG CTT TCA AGC ATC AAG CAC GAC CTC G 3'	*Hind*III	Tm= 62.9°C
oRD55(R): 5' CAT ATG TCC GCT TCC CCT ACA GCC 3'	*Nde*I	Tm= 61.3°C
oRD56(F): 5' GGA CAC GTC AGC AGC ATA TTT TGG TGC G 3'		Tm= 63.5°C

Table 2.3 - Gene sequence of XFLM_07435 from *X. fastidiosa* subsp. *fastidiosa* (ICMP 15197) as determined by sequencing, and translated amino acid sequence.

```
  1 ATGTCCGCTTCCCCTACAGCCGAGCGCTGGATCGTTCTAAAATTTGGTGGTACTTCCGTA
 +1 M  S  A  S  P  T  A  E  R  W  I  V  L  K  F  G  G  T  S  V
 61 TCGCGTCGTCATCGCTGGGACACCATTGCGATGGTGGTGAGGAAGCGCCTTGAAGAGCAT
 +1 S  R  R  H  R  W  D  T  I  A  M  V  V  R  K  R  L  E  E  H
121 GGAACACGTGTACTGATCGTGGTTTCGGCGCTGTCTGGCGTGACCAATGAACTGACGGCG
 +1 G  T  R  V  L  I  V  V  S  A  L  S  G  V  T  N  E  L  T  A
181 ATTGCTCAAGGTGTTGTGGACAGTGCCAAGCGCGTCACTGCATTGGAGCAGCGTCATCGC
 +1 I  A  Q  G  V  V  D  S  A  K  R  V  T  A  L  E  Q  R  H  R
241 GATTTTTTGGTTGAGCTTGGGTTAGATGCGCAAGCTGTGCTCGGCACGCGTTTCGCCATG
 +1 D  F  L  V  E  L  G  L  D  A  Q  A  V  L  G  T  R  F  A  M
301 TTGTCGGATCTTCTGCAAGATGCGCGTGCGGTGACGCGGTCACTGGACTGGCAGGCTGAG
 +1 L  S  D  L  L  Q  D  A  R  A  V  T  R  S  L  D  W  Q  A  E
361 TTGTTAGGGCAAGGTGAACTGTTGTCTTCCACACTCGGTGCGGCCTACCTGGGTGCTTCG
 +1 L  L  G  Q  G  E  L  L  S  S  T  L  G  A  A  Y  L  G  A  S
```

```
 421 GGTATTGATATCGGTTGGATGGATGCACGTGACTGGCTGACCGCTTTACCACCACAGCCC
  +1 G  I  D  I  G  W  M  D  A  R  D  W  L  T  A  L  P  P  Q  P
 481 AACCAGAGTGAATGGTCGCAACGTCTTTCTGTCTCCTGTCAATGGAAGTCTGATGTGGAG
  +1 N  Q  S  E  W  S  Q  R  L  S  V  S  C  Q  W  K  S  D  V  E
 541 TGGCGTACACGTTTCGATGCCCAACGTGCTCAGGTATTGATCACTCAAGGGTTTATCTCA
  +1 W  R  T  R  F  D  A  Q  R  A  Q  V  L  I  T  Q  G  F  I  S
 601 CGTCATCAAGATGGTGGTACCGCCATCCTGGGACGTGGTGGTTCGGACACGTCAGCAGCA
  +1 R  H  Q  D  G  G  T  A  I  L  G  R  G  G  S  D  T  S  A  A
 661 TATTTTGGTGCGCTACTTGGTGCGGCGTGTGTTGAGATCTGGACTGATGTACCAGGGATG
  +1 Y  F  G  A  L  L  G  A  A  C  V  E  I  W  T  D  V  P  G  M
 721 TTCAGTGCGAACCCGAAGGAGGTCCCTGATGCGCGTTTGCTGACGCGCCTGGATTACTAC
  +1 F  S  A  N  P  K  E  V  P  D  A  R  L  L  T  R  L  D  Y  Y
 781 GAGGCACAGGAGATCGCCACAACTGGCGCGAAGGTGTTGCATCCGCGTTCGATCAAACCA
  +1 E  A  Q  E  I  A  T  T  G  A  K  V  L  H  P  R  S  I  K  P
 841 TGCCGTGATACTGGCGTGCCGATGCTGATTCTAGATACTGAGCGGCCCGATGTGTTGGGG
  +1 C  R  D  T  G  V  P  M  L  I  L  D  T  E  R  P  D  V  L  G
 901 ACCCGTATTGATGCTGAAGTTGAGCCGGTGCTTGGAGTCAAGGCGATTAGCCGTCGTGAT
  +1 T  R  I  D  A  E  V  E  P  V  L  G  V  K  A  I  S  R  R  D
 961 GGCATTGTATTAGTGTCGATGGAAGGGATCGGCATGTGGCAGCAGGTGGGTTTTCTTGCT
  +1 G  I  V  L  V  S  M  E  G  I  G  M  W  Q  Q  V  G  F  L  A
1021 GACGTATTTACATTGTTCAAGAAACACGGATTGTCGGTGGACCTCATCGGATCGGCAGAA
  +1 D  V  F  T  L  F  K  K  H  G  L  S  V  D  L  I  G  S  A  E
1081 ACCAATGTTACTGTGTCCTTGGACCCCTCCGAGAACCTGGTTAACACCGACGTACTGACT
  +1 T  N  V  T  V  S  L  D  P  S  E  N  L  V  N  T  D  V  L  T
1141 GCATTGTCAACCGATTTATCACAGATCTGCAAAGTCAAAATTATTGTGCCTTGCGCTGCG
  +1 A  L  S  T  D  L  S  Q  I  C  K  V  K  I  I  V  P  C  A  A
1201 ATCACTTTGGTCGGGCGTGGTATGCGTTCCCTGCTGCATAAACTCTCTGAGGTATGGGCT
  +1 I  T  L  V  G  R  G  M  R  S  L  L  H  K  L  S  E  V  W  A
1261 ACGTTCGGTAAAGAACGTGTCCACATGATCTCGCAGTCTTCCAATGATCTTAATCTCACC
  +1 T  F  G  K  E  R  V  H  M  I  S  Q  S  S  N  D  L  N  L  T
1321 TTTGTGATTGATGAGACTGATGCGGACGGATTATTGGCGGTGCTGCATTTCAAATTGATC
  +1 F  V  I  D  E  T  D  A  D  G  L  L  A  V  L  H  F  K  L  I
1381 GAAAGTGGTGCAATGCCTGTCCAAGATCGTGCCGTGTTTGGTTCGCCTTGGCGCGAGATC
  +1 E  S  G  A  M  P  V  Q  D  R  A  V  F  G  S  P  W  R  E  I
1441 ATCGGGCACGTACGTACGCGGGTCATGCCTTGGTGGCACAGTGCCCGCATGCGTTTACTG
  +1 I  G  H  V  R  T  R  V  M  P  W  W  H  S  A  R  M  R  L  L
1501 GCCATGAGAGATGAAGAGGGCATTGGTCCTTGCTACGTCTACCATTTGCCGACAGTGCGC
  +1 A  M  R  D  E  E  G  I  G  P  C  Y  V  Y  H  L  P  T  V  R
1561 GAACGTGCGCGGCAATTGAATGCAGTAGCTGCCGTGGATCAGTGCTACTACGCCATCAAG
  +1 E  R  A  R  Q  L  N  A  V  A  A  V  D  Q  C  Y  Y  A  I  K
1621 GCCAACCCGCATCCGGCCATTCTCCGTACCCTGATTGAAGAAGGCTTTGGACTAGAGTGC
  +1 A  N  P  H  P  A  I  L  R  T  L  I  E  E  G  F  G  L  E  C
1681 GTGTCATTGGGTGAATTGCGTCACGTGTTTGAAGTGGTGCCACACTTGGTTCCGTCGCGT
  +1 V  S  L  G  E  L  R  H  V  F  E  V  V  P  H  L  V  P  S  R
1741 GTGTTATTCACTCCAAGTTTCGCCTCACGTGTGGAGTACGAGCAGGCATTCTTACTAGGT
  +1 V  L  F  T  P  S  F  A  S  R  V  E  Y  E  Q  A  F  L  L  G
```

```
1801 GTCAATGTGACCTTGGACAATGTCGAGGCCTTGCGCTTATGGCCGGAGGTATTCCGTGAA
  +1 V  N  V  T  L  D  N  V  E  A  L  R  L  W  P  E  V  F  R  E
1861 CGGACGTTATGGTTACGTATCGATTTGGGTTACGGCGATGGACATCACGAGAAGGTCACC
  +1 R  T  L  W  L  R  I  D  L  G  Y  G  D  G  H  H  E  K  V  T
1921 ACCGGTGGTAAGACATCTAAATTTGGCCTTTCGGTCACGCGGATTGATGAATTTATCGAT
  +1 T  G  G  K  T  S  K  F  G  L  S  V  T  R  I  D  E  F  I  D
1981 GTTGCCAAGACACTGAACATAACCGTGACCGGCCTGCATGCCCATCTTGGCAGTGGCGTG
  +1 V  A  K  T  L  N  I  T  V  T  G  L  H  A  H  L  G  S  G  V
2041 GAGAGCGGCAACCATTGGTCTCGCATGTACGATGAGTTGGCTGGTTTTGCCTCCCGCATC
  +1 E  S  G  N  H  W  S  R  M  Y  D  E  L  A  G  F  A  S  R  I
2101 GGGAGTGTGTCTACCCTGGATATCGGCGGCGGCTTACCGATCCCGTACCGGCCTGATGAT
  +1 G  S  V  S  T  L  D  I  G  G  G  L  P  I  P  Y  R  P  D  D
2161 GAACCATTTGACGTGGGGGCGTGGGGAGCAAGTTTGGCTGAAGTCAAAGCCTTGCATCCC
  +1 E  P  F  D  V  G  A  W  G  A  S  L  A  E  V  K  A  L  H  P
2221 AAATTTCAATTGGTGATTGAGCCGGGCCGCTTCCTTGTAGCCGAAGCGGGCGTGCTCTTG
  +1 K  F  Q  L  V  I  E  P  G  R  F  L  V  A  E  A  G  V  L  L
2281 ACCAGGGTCACCCAAGTGATTGAAAAAGATGACATCCTTCGGGTTGGCCTGGAGGCTGGC
  +1 T  R  V  T  Q  V  I  E  K  D  D  I  L  R  V  G  L  E  A  G
2341 ATGCATACCCTGATTCGTCCGGCCCTTTACGATGCCTGGCACGGCGTGGAGAACCTCACC
  +1 M  H  T  L  I  R  P  A  L  Y  D  A  W  H  G  V  E  N  L  T
2401 CGCTTGGATGAGACACCCCATGTACTGTGCGATGTCGTGGGTCCTATTTGCGAGTCATCC
  +1 R  L  D  E  T  P  H  V  L  C  D  V  V  G  P  I  C  E  S  S
2461 GACGTGTTTGGCCGTCGCCGTCACCTTCCAGCCGCCACCGCCTCCGGTGATTTAATGTTG
  +1 D  V  F  G  R  R  R  H  L  P  A  A  T  A  S  G  D  L  M  L
2521 ATTGCAGATGCTGGCGCCTATGGTTTCTCGATGGCCAGCACTTACAATCTGCGCGGCCTG
  +1 I  A  D  A  G  A  Y  G  F  S  M  A  S  T  Y  N  L  R  G  L
2581 CCGGCCGAGGTCGTGCTTGATGCCTGA
  +1 P  A  E  V  V  L  D  A  *
```

2.2.2 Agar Plates

Luria-Bertani (LB) broth base (Invitrogen™, by Life Technologies™, Carlsbad, California, USA) (20 g/L) was added to agar (15 g/L) and mixed thoroughly. The LB-agar mix was autoclaved and left to cool to approximately 50°C before adding the relevant antibiotic. Autoclaved LB-agar was stored at room temperature and melted in a microwave when needed. Molten LB-agar was poured into sterile petri dishes in a laminar flow hood.

2.2.3 Competent Cells

Competent *E. coli* cells stored at -80°C were; taken out of the freezer and a loop of cells was streaked out in a quadrant pattern on an LB agar plate containing no antibiotics. This plate was incubated at 37°C overnight. Single colonies were added to 3 mL pre-cultures (LB with no antibiotics) and placed on a shaker in a 37°C room overnight. A fraction of the pre-cultures (500 μL) was added to 25 mL of LB. These larger cultures were incubated at 37°C. When the O.D.600 reached 0.6–0.8, the cultures were put on ice. Then they were centrifuged at 3,000 rpm for five minutes at 4°C. The media was discarded and the cells were resuspended in 30 mL of ice cold calcium chloride (132 mM). The cells were incubated on ice for 30 minutes, then centrifuged at 3,000 rpm for five minutes at 4°C and the supernatant removed. The cells were resuspended in 2 mL of calcium chloride and 1 mL of 50% glycerol. Aliquots of 100 μL were pipetted into eppendorfs and flash frozen in liquid nitrogen. The cells were stored at -80°C. To test that the cells contained no contaminating plasmids, newly made competent cells were plated on LB agar containing ampicillin, kanamycin or no antibiotics. Cells that grew on the plates containing antibiotics were rejected as being contaminated. If the competent cells grew only on the plate without antibiotics they were considered satisfactory.

2.2.4 Polymerase chain reaction

Amplification of the AK:DapDc gene (XFLM_07435, Table 2.3) was carried out using Platinum® *Taq* DNA Polymerase High Fidelity (Invitrogen™ by Life Technologies™, Carlsbad, California, USA). The 10x High Fidelity PCR Buffer, 50 mM magnesium sulphate, and 50 mM dNTP mix was supplied with the Platinum® *Taq* DNA Polymerase. PCR reactions were run on a Multigene™ PCR Thermal Cycler (LabNet International Inc., Edison, New Jersey, USA).

For the initial amplification of the XFLM_07435 gene, 10 µL of 25 ng/µL genomic DNA was used. For most PCR reactions, 100 ng of DNA was used and made up to a volume of 10 µL with water. The reaction mix in Table 2.4 was used. Table 2.5 shows the general template used for the reactions. The annealing temperature was originally set at 55°C, but was decreased for subsequent PCR reactions due to lower melting temperatures of the primers (Table 2.2). Extension at 72°C was increased to two minutes due to the length of XFLM_07435 (2,604 base pairs).

Table 2.4 - PCR reaction parameters.

PCR Mix	**Volume**
10x buffer	5 µL
dNTPs (50 mM)	1 µL
Forward primer (100 µM)	0.2 µL
Reverse primer (100 µM)	0.2 µL
DNA (genomic or plasmid)	~100 ng in 10 µL
$MgSO_4$ (50 mM)	2 µL
Platinum® *Taq* polymerase Hi fidelity	0.2 µL
H_2O	30 µL
Total	**50 µL**

Table 2.5 - Template PCR Program

Temperature (°C)	Time (min)	Cycles
94°C (melting)	2	1
94°C (melting) ~3°C lower than primer Tm (annealing) 72°C (extension)	0.75 0.75 2	25
72°C (extension)	2	1
4°C	As needed (<60)	1

2.2.5 DNA Gel Electrophoresis

All DNA gels were made to a concentration of 1% agarose with 1xTris-acetate-EDTA (TAE) buffer, diluted from a 50x TAE stock solution. The 50x TAE stock solution contained: 242 g Tris Base, 57.1 mL glacial acetic acid, and 100 mL 0.5 M EDTA pH 8.0, made up to a volume of 1 L with H_2O.

Agarose (0.5 g) was added to 50 mL TAE, which was then heated until all the agarose was dissolved. SYBR Safe (Invitrogen™ by Life Technologies, Carlsbad, California, USA) (3 µL) was added to the heated solution and mixed. PCR product (20 µL) was added to 5-10 µL of loading buffer and run against a DNA ladder (HyperLadder I, Bioline USA Inc., Taunton, Massachusetts, USA) at 120 V for 20 minutes. Gels were visualised with a Syngene ‘Gene Genius’ gel doc (Syngene UK, Cambridge, UK) and GeneSnap software.

2.2.6 Ligation into Vector

When ligating into linearised pCR2.1-TOPO vector, 2 μL of PCR product, 1 μL of vector (1 μL = 1 reaction), 1 μL of salt solution and 2 μL of water were incubated at room temperature for 20 minutes.

For ligation of XFLM_07435 into linearised pBluescriptSK+, pET30GST or pET30ΔSE, a 2:1 ratio of insert to plasmid was used (approximately 20 ng of insert to 10 ng of vector) while trying to keep the reaction volume less than 10 μL. To the combined insert and plasmid, 6 μL of T4 DNA ligase (DNA Ligation Kit Ver.2.1, TaKaRa Bio Inc., Otsu, Shiga, Japan) was added. The reaction was incubated at room temperature for 20 minutes.

2.2.7 Transformation

Ligated plasmid (100 ng) was added to a tube of competent cells (100 μL) on ice. Competent cells with no DNA added were used as a control. The cells were incubated on ice for 20 minutes, heat shocked for 90 seconds at 42°C, and then returned to ice for another two minutes. LB (200 μL) was added to the cells for recovery. The transformed cells were then incubated at 37°C for one hour before being plated on agar with the relevant antibiotic. The plate was incubated at 37°C overnight.

2.2.8 Plasmid Extraction

Plasmid extraction was achieved according to the protocol in the Purelink® Quick Plasmid Miniprep Kit (Invitrogen™ by Life Technologies, Carlsbad, California). A 5 mL pre-culture of *E. coli* cells containing the desired plasmid was grown overnight at 37°C. The cells were centrifuged at 12,000 rpm for 10 minutes and the resulting pellet was used for the miniprep. Cells were first resuspended with the resuspension buffer (250 μL), then lysed with 250 μL of lysis buffer, at room temperature for five minutes. Precipitation buffer

(350 μL) was added to the lysate. The lysate was centrifuged at 12,000 rpm for 10 minutes to remove cell debris. The supernatant was loaded into a spin column and centrifuged at 12,000 rpm for one minute. Wash buffer (700 μL) was applied to the spin column and centrifuged for one minute at 12,000 rpm. The flowthrough was discarded and the column was spun for one minute to remove the wash buffer. The spin column was placed in a recovery tube and 50 μL of H_20 was added to the column. After incubation for one minute, the column was centrifuged at 12,000 rpm for two minutes to elute the DNA. The DNA concentration was measured at 280 nm using a NanoDrop Spectrophotometer ND-1000 (Thermo Scientific, Waltham, Massachusetts, USA). Purified DNA was stored at -20°C.

2.2.9 Restriction Digest

For each sample to be digested, 1000 ng of DNA was required. In all digestions, 5 μL of 10xNEBuffer #4 (500 mM potassium acetate, 200 mM Tris-acetate, 100 mM magnesium acetate, 10 mM DTT, pH 7.9), 1 μL of each restriction enzyme (20,000 U/mL), and 0.2 μL of BSA (10 mg/mL) were added. The reaction mix was brought up to 50 μL with water. The digest was incubated at 37°C for one-three hours. The digested samples (50 μL) were then loaded onto an agarose DNA gel and run for 20 minutes. The DNA bands were visualised using a DarkReader transilluminator. DNA bands were cut out of the agarose gel and placed in 2 mL eppendorf tubes.

2.2.10 DNA Purification from Agarose Gel

DNA was purified from agarose gel using an Agarose Gel DNA Extraction kit (Roche, Indiana, Indianapolis, USA) and a small benchtop centrifuge. Agarose solubilisation buffer (500 μL) was added to a piece of agarose gel containing the DNA of interest (from section 2.2.9). This was heated to 55°C for five minutes to melt the agarose. 10 μL of silica matrix

was added to bind the DNA. It was mixed thoroughly and incubated at room temperature for 10 minutes. The silica was pelleted (12,000 rpm for 30 seconds) and the supernatant discarded carefully so as not to remove any silica. 500 μL of nucleic acid binding buffer was added and mixed with the silica. The mixture was centrifuged, and the supernatant discarded. 500 μL of washing buffer was added to the silica and mixed, and then the sample was centrifuged for one minute. All of the supernatant liquid was then removed using a pipette. The silica was left to dry for a few minutes. The silica was then resuspended in 10 μL of nuclease-free water and vortexed every two-three minutes for 10 minutes. Centrifugation at 12,000 rpm for 30 seconds pelleted the silica, and the supernatant containing the purified DNA was removed. The concentration of DNA was measured using the Nano Drop.

2.2.11 DNA Sequencing

Sequencing of the XFLM_07435 gene was performed in the sequencing suite at the University of Canterbury. Three primers were used to sequence the entire gene (2.2.1.4), as XFLM_07435 is a long gene (2,604 base pairs). The primers were used at a concentration of 3.2 μM. Miniprepped dsDNA plasmid was provided at a concentration of 50 ng/μL, with 6-7 μL given per primer used in sequencing (e.g. 3 primers = 20 μL).

2.3 Protein Biochemistry

2.3.1 Pre-cultures

To make pre-cultures, 5 µL of relevant antibiotic and one colony or a stab from a frozen glycerol stock were added to 5 mL of autoclaved LB. The pre-culture was incubated at 37°C overnight on a shaker.

2.3.2 Glycerol Stocks

A 5 mL pre-culture was incubated overnight at 37°C. 1 mL aliquots were pipetted into 2 mL eppendorf tubes. 500 µL of 50% glycerol was added to each aliquot of culture. The cells were then snap frozen using liquid nitrogen and stored at -80°C.

2.3.3 Protein Expression

A pre-culture (2-5 mL) was added to 500 mL of autoclaved LB containing the appropriate concentration of antibiotic. These cultures were placed in a Minitron incubator at 37°C and grown to an OD_{600} of 0.4-0.6 for cells transformed with the pET30ΔSE-HisXF vector or to an OD_{600} of 0.7-0.9 for cells transformed with SDH-pET16b. Cultures were then incubated at 26°C for 30 minutes to cool down before inducing with IPTG. Cells containing XFLM_07435 were induced with 0.2 mM IPTG. Cells containing SDH were induced with 1 mM IPTG. The cells were then incubated overnight at 26°C to allow expression. Cells were harvested by centrifugation at 8,000 rpm for 15 minutes. Cell pellets were stored at -20°C or put on ice and resuspended with lysis buffer.

2.3.3.1 Expression Trials

Expression trials were conducted using small scale cultures (20-100 mL LB). Cells transformed with a vector containing the XFLM_07435 insert (pBluescript-XF,

pET30ΔSE-XF, pET30ΔSE-HisXF, or pET30GST-XF) were plated on agar containing the relevant antibiotic(s). The agar plate was incubated overnight at 37°C. The next day, 5 mL pre-cultures were made from individual colonies. These pre-cultures were incubated at 37°C overnight. The next day, each pre-culture was added to a flask of autoclave LB (20-100 mL), with the relevant antibiotic(s), and the culture was incubated at 37°C until the OD_{600} reached 0.4-0.6. The cultures were then induced with 1 mM IPTG (final concentration) and expressed overnight at 37°C. After samples were taken to test for expression or solubility the remainder of the culture was centrifuged at 8,000 rpm for 10 minutes at 4°C and the media removed. Cell pellets were stored at -20°C.

2.3.3.2 Testing for over-expression

Samples (20 μL) were taken pre-induction, one hour post-induction, four hours post-induction and 16 hours post-induction. These samples were analysed by SDS-PAGE (Section 2.3.4) for over-expression. The remainder of the cultures were centrifuged at 8,000 rpm, then stored at -20°C to be used for subsequent solubility tests.

2.3.3.3 Testing for solubility

Larger samples were taken (1 mL-20 mL) pre-induction, one hour post-induction, four hours post-induction, and 16 hours post-induction. These samples were centrifuged at 8,000 rpm and the media removed. The cell pellets were resuspended in buffer: 1 mL of culture was resuspended in 100 μL of buffer. The samples resuspended in 100-200 μL buffer was lysed using a sonicating bath for 30 second intervals for a total of two minutes of sonicating. Samples resuspended in 200 μL or more of buffer were lysed using the Stepped Titanium Microtip (3.9 mm diameter) of the Sonic Ruptor 400 Ultrasonic Homogenizer (Omni International Inc., Kennesaw, Georgia, USA). Sonication was carried

out on ice for 5 second intervals for a total of two minutes. Lysed samples were then centrifuged for 10 minutes at 8,000 rpm and 4℃. The supernatants were analysed by SDS-PAGE (Section 2.3.4).

2.3.4 Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis

Bolt® Pre-cast 4-12% Bis-Tris Plus 10 well or 15 well electrophoresis gels were supplied by Novex (Novex® by Life Technologies™, Carlsbad, California, USA). Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS PAGE) was carried out in a Bolt™ Mini Gel Tank (Novex® by Life Technologies™, Carlsbad, California, USA) using a BioRad PowerPac machine with 1x MES SDS buffer (Diluted from 20x MES SDS Buffer, Invitrogen™ by Life Technologies, Carlsbad, California, USA). Novex® Sharp Pre-stained Protein Standard (Novex® by Life Technologies, Carlsbad, California, USA) was used as a ladder/reference by loading 5 µL into one of the wells of the gel. Sample reducing dye (25 µL Bolt™ LDS Sample buffer 4x, 10 µL Bolt™ reducing agent 10x, and 15 µL H_2O) was added to the samples at a ratio of 1:1. Samples were then boiled at 90℃ in an Accublock Digital Dry Bath (LabNet International Inc., Edison, New Jersey, USA) for two minutes before loading them into the gel. Boiled samples were stored at -20℃ if not used immediately. The gels were run at 165 V for 35 minutes. Gels were stained with Coomassie Blue stain (1 g Coomassie, 450 mL methanol, 450 mL water, 100 mL acetic acid) for 20 minutes and destained using Coomassie destain (450 mL methanol, 450 mL water, 100 mL acetic acid) until the gel was destained sufficiently to photograph the gel using the Syngene ‘Gene Genius’ gel doc (Syngene UK, Cambridge, UK).

2.3.5 Cell Lysis

Cell pellets from 1 L cultures containing AK:DapDc were resuspended in 10 mL lysis buffer I (20 mM Tris pH 8.5, 200 mM NaCl, 1 mM DTT, 5 μM PLP) and lysed using a Microfluidics M-110P homogenizer (Microfluidics USA, Westwood, Massachusetts, USA). The machine was first washed with H_2O to remove all methanol from the system and then equilibrated with the cell lysis buffer. The resuspended cells were run through the machine two-three times to ensure complete lysis, according to the manual (Microfluidics M-110P series). The lysate was then centrifuged for 10 minutes at 9,000 rpm.

Cell pellets containing *Sc*SDH were resuspended in lysis buffer II (200 mM Tris pH 8.0) and lysed using a Hielscher UPS200 S ultrasonic processor, according to Xu *et al*. (2006)[1]. The sonicator was set to a cycle of 0.3 second pulse followed by a 0.7 second rest at 40% power for two minutes. The lysate was then centrifuged at 9,000 rpm for 10 minutes.

2.3.6 Column Chromatography

2.3.6.1 HisTrap Chromatography of AK:DapDc

A 5 mL pre-packed HiTrap FF HisTrap column (GE Healthcare Life Sciences, Uppsala, Sweden) was used for AK:DapDc purification. The HiTrap FF column was controlled by an ÄKTAFPLC (GE Healthcare, Uppsala, Sweden) and fractionation was carried out by the same machine. The 5 mL column was pre-equilibrated with five column volumes of lysis buffer I (See below). The lysed and centrifuged crude fraction (40-80 mL) was loaded onto the column at 2 mL/minute. Once the entire sample was loaded, the column was washed with 40 mL of lysis buffer I at 2.5 mL/minute. A gradient of 0-100% of buffer B (Table 2.6) over 30 minutes was initiated. The ÄKTA-FPLC was set to run for an

additional 10 minutes at 100% buffer B. Fractions of 1.8 mL were collected throughout the purification, beginning as the centrifuged crude fraction was loaded onto the column.

Table 2.6 - Purification buffers.

Buffer	Composition
A	Lysis buffer I (20 mM Tris pH 8.5, 200 mM NaCl, 1 mM DTT, 5 μM PLP)
B	Lysis buffer I + 500 mM imidazole

2.3.6.2 His-Trap Chromatography of ScSDH

His-Trap chromatography was carried out using a 5 mL disposable gravity flow column (Thermo Fisher Scientific Inc., Rockford, Illinois, USA) and 1 mL of HisPur™ Cobalt Resin (Thermo Fisher Scientific Inc., Rockford, Illinois, USA) for SDH purification. The HisPur Cobalt column was operated on the bench top and 1 mL fractions were collected in 1.7 mL eppendorf tubes. The 1 mL column was pre-equilibrated with 10 mL of lysis buffer II (200 mM Tris pH 8.0). The lysed and centrifuged crude fraction (20-40 mL) was loaded onto the column using a syringe at approximately 3 mL/minute. The flowthrough was collected in a 50 mL Falcon tube. The column was then washed with 20 mL of lysis buffer II, followed by 20 mL of lysis buffer II + 20 mM imidazole. The flowthrough from both washes was also collected in a 50 mL Falcon tube. SDH was eluted from the column using 20 mL of lysis buffer II + 150 mM imidazole. Elution fractions of 1 mL were collected in 1.7 mL Eppendorf tubes. Samples (10 μL) of the lysed cells, crude fraction, flowthrough, both washes, and elution fractions were analysed using SDS-PAGE.

2.3.6.3 Size Exclusion Chromatography (SEC)

SEC was carried out using a HiLoad 16/200 Superdex 200 pg (GE Healthcare Life Sciences, Uppsala, Sweden) column with a volume of 120 mL. This column was also controlled by an ÄKTA and fractions of 1 mL were collected by the same machine throughout the purification. The column was equilibrated with 400 mL of SEC buffer (20 mM Tris pH 7.5, 200 mM NaCl, 1 mM DTT, 5 μM PLP). The protein sample from His-Trap chromatography was concentrated to 2 mL using a Vivaspin® 6 30,000 Da MWCO spin concentrator (Sartorius, Göttingen, Germany). The sample was loaded into a 2 mL loop and loaded onto the column at a rate of 1 mL/minute. Once the sample was loaded, 120 mL of SEC buffer was run through the column at 0.5 mL/minute. Fractions of the elution peaks were analysed by SDS-PAGE.

2.3.7 Kinetic Analysis

2.3.7.1 Stock Solutions

Tris buffer (pH 8.0) was made as a stock solution of 1 M in H_2O. DAP is not very soluble, so it was dissolved in equimolar Tris, that is 200 mM DAP was dissolved in 200 mM Tris pH 8.0. Lysine was dissolved in H_2O to give a stock solution of 500 mM. PLP stock solution was made up in H_2O to a concentration of 5 mM. α-ketoglutarate (α-KG) is very acidic and was made as a stock solution of 500 mM by dissolving the α-KG in half the final stock volume, then using 5 M NaOH to increase the pH to 8.0, before making up the volume to the final stock volume. The SDH stock solution concentration was 1.5 mg/mL and was achieved by concentrating the SDH elution fractions from HisTrap chromatography using a Vivaspin 6 10,000 Da MWCO spin concentrator (Sartorius, Göttingen, Germany).

2.3.7.2 Equipment and Software

Kinetic analysis was carried out using a CARY 100 Bio UV-Visible Spectrophotometer (Agilent Technologies, Santa Clara, California, USA), using the CARY Kinetics software.

2.3.7.3 Assay reaction mix

The SDH activity assay was carried out with the assay mix listed in Table 2.7. The total volume of the assay mix was actually 1000 µL minus the volume of SDH to be used in the assay, so that the total volume of every assay is 1000 µL.

Table 2.7 - The saccharopine dehydrogenase assay reaction mix.

Chemical	**MW**	**Initial (mM)**	**Volume (µL)**	**Final (mM)**
Tris pH 8.0	121.14	1000	135	135
DAP/Tris	190.2	200/200	200	40/40
Lysine	182.6	500	50	25
PLP	247.142	5	20	0.1
α-KG/Tris	146.11	500/500	50	25/25
NADH (Na_2)	709.4	8	20	0.16
Water			up to 1000	
Total			**1000**	

2.3.7.4 SDH concentration vs. rate of reaction assay

The assay was monitored at 340 nm for two minutes at 37°C. The instrument was blanked with water. The initial slope of the reaction was estimated using the CARY Kinetics software. The reaction mix was set up as per Section 2.3.7.3, using an SDH stock of 1.5 mg/mL (34.1 µM). The final concentration of SDH in the 1 mL reaction was increased

from 0.0335 μM to 1.5 μM, with the volume of water added adjusted so that the final reaction volume was 1 mL. The SDH concentrations measured were: 0.0335 μM, 0.0675 μM, 0.125 μM, 0.25 μM, 0.5 μM, 1.0 μM, 1.5 μM. No measurements were taken above 1.5 μM SDH as the initial rate became too difficult to measure and calculate accurately. Three replicates of each concentration were measured. These replicates agreed within 10% error.

2.3.7.5 Lysine concentration vs. rate of reaction assay

The assay was monitored at 340 nm for two minutes at 37°C. The instrument was blanked with water. The initial slope of the reaction was estimated using the CARY Kinetics software. The reaction mix was set up as per Section 2.3.7.3, using an SDH stock of 1.5 mg/mL (34.1 μM) and the final concentration of SDH in the reaction was 0.25 μM. The concentration of lysine ranged from 1-100 mM (1 mM, 2 mM, 5 mM, 10 mM, 15 mM, 25 mM, 50 mM, 100 mM) while keeping the α-ketoglutarate(α-KG) concentration at 25 mM. The volume of water added was adjusted so that the final reaction volume was 1 mL.

2.3.7.6 Assaying for AK:DapDc activity

The assay was monitored at 340 nm for two minutes at 37°C. The instrument was blanked with water. The initial slope of the reaction was estimated using the CARY Kinetics software. The reaction mix was set up as per Section 2.3.7.3, using an SDH stock of 1.5 mg/mL (34.1 μM) and a final concentration of SDH in the reaction of 3 μM. Lysine was excluded from the reaction mix as this assay tests the ability of AK:DapDc to convert *meso*-diaminopimelate into lysine. The AK:DapDc stocks used were 1 mg/mL (10.25 μM). The final concentration of AK:DapDc used was 200 nM. For this assay, the reaction mix

was equilibrated to 37ºC, then the SDH enzyme was added and mixed using a pipette. The absorbance was monitored for a few seconds to check that the SDH was not reacting with the substrates present. The AK:DapDc was then added and mixed quickly with a pipette. The absorbance was monitored for two minutes.

2.4 Biophysical Techniques

2.4.1 Differential Scanning Fluorimetry (DSF)

DSF was carried out on a Bio-Rad iQ5 RT-PCR machine (Bio-Rad Laboratories, Hercules, California, USA), using Bio-Rad iQ5 software. 96 well clear pcr plates and clear seals were used. The fluorescence of SYPRO Orange was measured from 20°C-100°C, in steps of 0.5°C. The temperature was increased for 10 seconds and held at that temperature for 10 seconds, taking measurements at each temperature.

Buffer optimisation using DSF was carried out using a modified Seabrook & Newman protocol [2], with the sample volume changed from 20 µL to 25 µL (0.3 µL of 1:10 SYPRO Orange, 2 µL protein sample, 22.7 µL buffer). See Appendix I for the modified table of buffers used. The 96 well plate was set up with the buffers, then protein sample, then SYPRO solution. Before sealing the plate with a clear seal the samples were degassed for 5 minutes.

For aggregation tests, a photo of the 96 well plate was taken to show the fluorescence detected at the start of the run. Buffer controls and positive controls (aggregated protein) were used.

2.4.2 Buffer Exchange

Buffer exchange of AK:DapDc samples was performed using a 5 mL GE HiTrap desalting column (GE Healthcare Life Sciences, Uppsala, Sweden) according to the GE Healthcare manual, with a syringe. The column was equilibrated with 25 mL of the final protein buffer. 1 mL of protein sample was applied to the column with a syringe. Buffer (0.5 mL) was added to the column to bring the total sample volume to 1.5 mL (the column's void volume). Additional buffer was applied to the column and the eluted buffer was collected until three 1 mL elution fractions had been eluted. A NanoDrop, measuring at 260 nm and 280 nm, was used to determine where the protein was eluted. The column was then washed and stored in 20% ethanol.

2.4.3 Circular Dichroism (CD)

Circular dichroism was carried out in a JASCO J-815 CD Spectrometer (JASCO International Co. Ltd, Hachioji, Tokyo, Japan) using a 2 mm cuvette. Scans were run from 180-260 nm (far-UV) at a pitch of 1 nm, with replicates of three. The speed was set to 100 nm/min. The temperature was set at 20℃. First, buffer checks were carried out to determine what concentration of Tris buffer and additives were acceptable to use for CD measurements. Samples were equilibrated at 20℃ for five minutes. Then measurements were taken from 180-260 nm.

2.4.4 Analytical Ultracentrifugation

Analytical ultracentrifugation to measure sedimentation velocity was performed on a Beckman Coulter Proteome Lab XL-I (Beckman Coulter, Brea, California, USA) ultracentrifuge, with an An-50 Ti 8-hole analytical rotor (Beckman Coulter), and using double sector cells with quartz windows. The density and viscosity of buffers was

calculated using Sednterp (http://sednterp.unh.edu/). Data analysis was performed using sedfit [3].

A 400 μL reference sample of buffer (SEC buffer) was placed in the reference cell. The sample volume was 380 μL. Three different concentrations of protein sample were run (0.1 mg/mL, 0.4 mg/mL, and 0.8 mg/mL). Initial wavelength (200-300 nm) and radial scans were conducted at 3,000 rpm to determine the best wavelength and radial position for the measurements. Runs were then carried out at 40,000 or 45,000 rpm, at a temperature of 20℃.

2.4.4.1 AK:DapDc sample purified using anion exchange and SEC

AK:DapDc (1 mg/mL) was divided into three 380 μL samples of 0.1 mg/mL, 0.4 mgmL and 0.8 mg/mL, adding SEC buffer to each sample to dilute them to the appropriate concentration. The concentration was measured using a spectrophotometer at 280 nm. Each of these samples (380 μL) was loaded into a double sector cell, with a buffer reference (SEC buffer) of 400 μL. The samples were spun at 3,000rpm, measuring at 200-300 nm to calculate the appropriate wavelength and radial position to take further measurements at. The sample was then run at 40,000 rpm, taking 250 scans at 290 nm and a radial position of 5.8-7.3 mm.

2.4.4.2 AK:DapDc sample purified by His-Trap chromatography and SEC

The elution peaks two and three containing AK:DapDc purified by His-Trap chromatography and SEC were pooled separately and concentrated to 1 mg/mL. Peak two was divided into three 380 μL samples of 1 mg/mL, 0.4 mg/mL, and 0.8 mg/mL, adding SEC buffer to each sample to dilute them to the appropriate concentration. The protein concentration was measured using a spectrophotometer at 280 nm. Each 380 μL sample

was loaded into a double sector cell, with a buffer reference (SEC buffer) of 400 μL. The samples were spun at 3,000 rpm, measuring at 200-300 nm to calculate the appropriate wavelength and radial position to take further measurements at. The samples were then run at 40,000 rpm, taking 150 scans at 288 nm and a radial position of 5.8-7.3 mm.

Peak three was divided into two samples of 0.1 mg/mL and 0.8 mg/mL, adding SEC buffer to each sample to dilute them to the appropriate concentration. The protein concentration was measured using a spectrophotometer at 280 nm. Each 380 μL sample was loaded into a double sector cell, with a buffer reference (SEC buffer) of 400 μL. The samples were spun at 3,000 rpm, measuring at 200-300 nm to calculate the appropriate wavelength and radial position to take further measurements at. The samples were then run at 45,000 rpm, taking 150 scans at 285 nm and a radial position of 5.8-7.3 mm.

2.4.5 Crystal Trays

Protein samples (AK:DapDc) were concentrated to 10 mg/mL using a Vivaspin 6 30,000 Da MWCO spin concentrator (Sartorius AG, Göttingen, Germany).

Morpheus™ (MD1-47), JCSG-*plus*™ HT-96 (MD1-36), PACT *premier*™ HT-96 (MD1-40), and Clear Strategy Screen™ I HT-96 (MD1-31) crystal screens from Molecular Dimensions (Molecular Dimensions Limited, Suffolk, UK) were used in this study. Crystallisation buffer from these screens (300 μL) was pipetted into the wells of a 96 well sitting drop plate using a multi-channel pipette. Two plates of each screen were pipetted out so that two crystallisation temperatures could be trialled- 5℃ and 20℃. A TTP LabTech Mosquito (TTP Labtech Ltd, Melbourn, Hertfordshire, UK) was used to pipette 300 nL of concentrated protein and 300 nL of crystallisation buffer onto the sitting drop well/plate. The trays were covered with clear seals. The plates were incubated at 5℃ or

20℃ for several weeks to see if crystals would form. Photos of the crystal trials were taken with a microscope camera.

References

1. Xu, H. Y., West, A. H., and Cook, P. F. (2006) Overall kinetic mechanism of saccharopine dehydrogenase from Saccharomyces cerevisiae, *Biochemistry 45*, 12156-12166.
2. Seabrook, S. A., and Newman, J. (2013) High-Throughput Thermal Scanning for Protein Stability: Making a Good Technique More Robust, *Acs Combinatorial Science 15*, 387-392.
3. Schuck, P. (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling, *Biophysical Journal 78*, 1606-1619.

3. Cloning, Expression and Purification of AK:DapDc

3.1 Introduction

Xylella fastidiosa research has primarily focused on its biology to establish how the bacterium infects and affects plants[1-5], which plants it affects[6-9], how it affects the insect vector, and which insects are vectors[10, 11]. Metabolic pathways such as carbohydrate metabolism have been investigated[12], but no research has focused on essential biosynthetic enzymes. Genome annotation of *X. fastidiosa*[13] suggested that a number of genes in biosynthetic pathways, such as amino acid biosynthesis, are missing[14]. Investigating the functionality of the predicted aspartate kinase:diaminopimelate decarboxylase, by first over-expressing and purifying it, could indicate if *X. fastidiosa* can synthesise lysine.

A bifunctional biosynthetic enzyme such as AK:DapDc, unique to a small family of bacteria (Xanthomonadaceae), might be a useful pesticide target, as an inhibitor could be highly specific to the bifunctional bacterial enzyme and not affect the corresponding monofunctional plant enzymes. However, AK:DapDc needs to be experimentally characterised first.

Characterisation of more bifunctional enzymes, especially those that catalyse non-consecutive reactions, will add to our knowledge base of the mechanisms of action of bifunctional enzymes. It may also help us to understand why these enzymes evolve and if they are more advantageous than multi-enzyme complexes. The evolutionary significance of enzymes such as AK:DapDc is that they may provide rapid regulation of pathways and colocalization of enzymes from the same pathway[15].

Many aspartokinases have been purified and crystallised, with the aspartokinases of *E.coli* being the most studied[16]. Aspartokinases I (bifunctional with HSDH I) and III originating from microorganisms such as *Escherichia coli* (BL21DE3, K12)[17], Anheuser Baker's yeast[18], *Streptomyces claviguligerus*[19], *Mycobacterium tuberculosis*[20], *Amycolatopsis mediterranei* U32[21], *Bacillus subtilis*[22], and *Synechocystis* species[23] have been purified. These studies[24] used endogenous genes to synthesise the protein. Aspartate kinase was purified by ammonium sulphate precipitation, ion exchange chromatography, and size exclusion[22, 25]. More recent papers replace ammonium sulphate precipitations with His-Trap chromatography and hydrophobic interaction chromatography[21], and use vectors to over-express recombinant genes in *E.coli*[20].

Diaminopimelate decarboxylase (DapDc) has been expressed and purified via several methods[26, 27]. Endogenous DapDc has been purified from microorganisms such as *E. coli*[26], *Micrococcus glutamicus*[28], *Corynebacterium glutamicum*[29], and *Bacillus subtilis*[30]. DapDc was purified by sonication/French press followed by ammonium sulphate precipitation, then size exclusion, anion exchange or hydroxyapatite chromatography. Recombinant DapDc from organisms such as *E. coli*[31], *Helicobacter pylori*[32], and *Mycobacterium tuberculosis*[33] have been over-expressed in *E. coli* and purified using Ni^{2+} His-Trap chromatography followed by size exclusion.

The functions of AK:DapDc are putative and it has never been experimentally characterized before[13]. Genome sequencing and annotation using BLAST[34] predict that AK:DapDc is one gene, containing the conserved domain features for AK and DapDc. In this study, information from the literature about the purification of the individual monofunctional forms of AK and DapDc has been used to attempt to find the best conditions with which to clone, express and purify the bifunctional AK:DapDc.

3.2 Cloning XFLM_07435 into an Expression Vector

The XFLM_07435 gene was amplified via polymerase chain reaction (PCR) from the genomic *X. fastidiosa* DNA to include the addition of the *Sac*I restriction site at the 5' end of the gene and the *Hind*III restriction site at the 3' end of the gene. PCR was carried out according to Section 2.2.1 using primers oRD53 (5' end) and oRD54 (3' end). This generated a 2604 bp sized DNA fragment, which gave a single band on a DNA agarose gel carried out according to Section 2.2.5 (Figure 3.1).

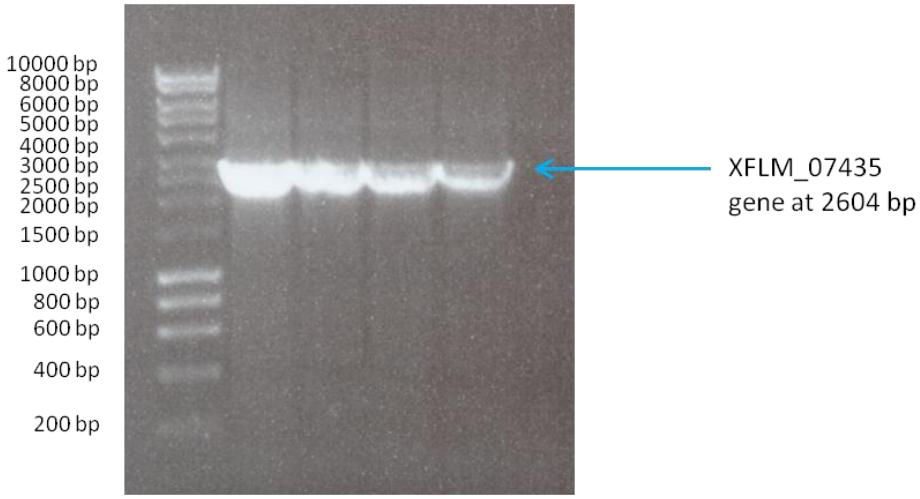


Figure 3.1 - PCR amplification of XFLM_07435.

Ligation of the PCR product into pCR2.1TOPO according to Section 2.2.6, followed by transformation of the ligation product into *E. coli* XL1 Blue cells (2.2.7), and plating on agar + Amp resulted in many transformed colonies. After pre-culturing several of the transformed colonies (2.3.1) and purifying the pCR2.1TOPO-XF plasmid from the cells (2.2.8), pCR2.1TOPO-XF and empty pET30ΔSE expression vector were subject to a double restriction digest using the *Sac*I and *Hind*III restrictions enzymes (2.2.9). pCR2.1TOPO-XF was digested to excise the XFLM_07435 gene from the plasmid ready for ligation into the expression vector pET30ΔSE. Both digested plasmids were run on an

agarose gel (Figure 3.2) so that the XFLM_07435 insert and the restricted pET30ΔSE could be purified according to Section 2.2.10 (pCR2.1TOPO + XFLM_07435 = 6,500 bp). Additional pET30ΔSE was digested with *Nde*I and *Hind*III to give a linearised vector with no His-Tag. pBluescriptSK+ and pET30GST were linearised by double digestion with *Sac*I and *Hind*III restriction enzymes.

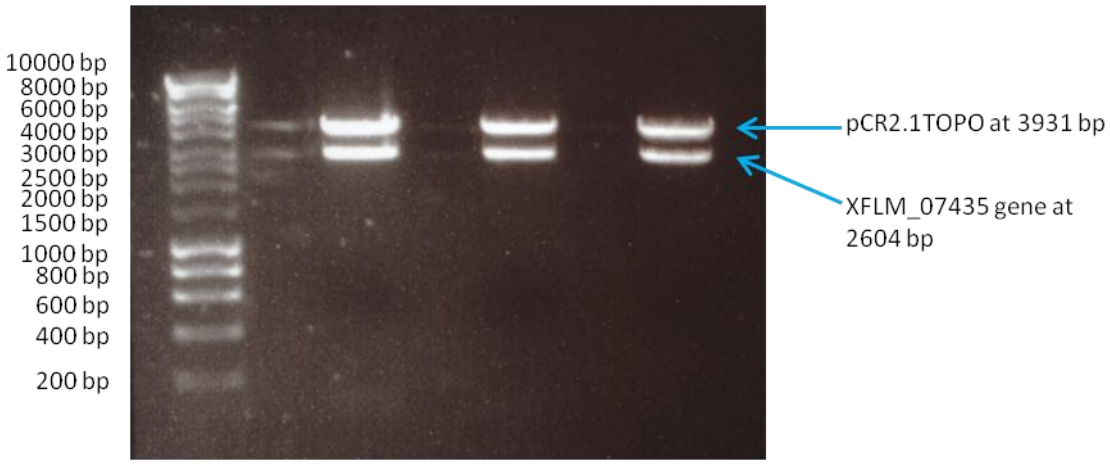


Figure 3.2 - DNA gel of the double digest of pCR2.1TOPO-XF to give pCR2.1TOPO and XFLM_07435 insert with 5' *Sac*I and 3' *Hind*III restriction sites.

The linearised pET30ΔSE vector and XFLM_07435, pBluescript SK+ and XFLM_07435, and pET30GST and XFLM_07435 were ligated according to Section 2.2.6 to give pET30ΔSE-XF, pET30ΔSE-HisXF, pBluescript-XF and pET30GST-XF. *E. coli* XL1 Blue cells were transformed with the ligated plasmids and incubated on LB-agar plates. Several resulting colonies from each transformation were pre-cultured, and plasmid DNA purified from the cells. This purified plasmid was sequenced using T7 vector primers, and this proved that the XFLM_07435 gene had been successfully inserted into the vector in each case. The pET30GST-XF construct could not be sequenced properly.

3.3 Expression

Over-expression of AK:DapDc was achieved after many expression trials (see section 2.3.3). *E. coli* BL21 DE3 cells were transformed with one of the purified vectors with XFLM_07435 inserts (pBluescript-XF, pET30GST-XF, pET30ΔSE-HisXF, or pET30ΔSE-XF). Small scale expression was trialled at 37°C. Cells transformed with pBluescript-XF did not over-express AK:DapDc, and despite sequencing results the cells transformed with pET30GST-XF only over-expressed the GST tag. The cells transformed with pET30ΔSE-HisXF or pET30ΔSE-XF over-expressed AK:DapDc at 37°C. Further expression trials were conducted with only the pET30ΔSE constructs.

The pET30ΔSE-HisXF was transformed into both *E. coli* BL21 DE3 and BL21 DE3 Rosetta cells and expression was carried out overnight at 37℃ to see if the rare tRNA codons encoded on the pRARE plasmid in Rosetta cells increased the solubility of AK:DapDc. There was no noticeable difference, so all subsequent expression trials were conducted using *E. coli* BL21 DE3 cells.

Several growth temperatures (4℃, 26℃, and 37°C) were trialled to determine the conditions best conducive to soluble protein expression. Expression at 26℃ gave the best AK:DapDc solubility (using the steps in section 2.3.3), so all subsequent trials were performed at 26°C. Low isopropyl β-D-1-thiogalactopyranoside (IPTG) concentration (0.1 mM) and expression overnight at 4℃, as well as high IPTG concentration (2 mM) and expression for two hours at 37℃ were tested. These conditions did not prevent AK:DapDc aggregation. Varying the IPTG concentration (0.1–1 mM) did not affect AK:DapDc solubility.

Supplementation of the LB media did not noticeably affect AK:DapDc expression or solubility. Glucose (1%) was added to both the pre-cultures and larger cultures to prevent leaky expression and slow down AK:DapDc expression. The cofactor for the DapDc domain (pyridoxal phosphate - PLP) was added to the media upon induction with IPTG to determine if an increase in PLP would help stabilize the DapDc domain. It remains unclear if this helps with solubility.

AK:DapDc was over-expressed in *E. coli* BL21 cells (Figure 3.3). Cultures of *E. coli* BL21 DE3 cells transformed with pET30ΔSE-HisXF were incubated at 37℃ until the $O.D_{600}$ reached 0.4–0.6, and then the cultures were incubated at 26℃ for 30 minutes to cool the cultures before inducing with 0.4 mM IPTG. Cultures were incubated overnight at 26℃. Samples of un-induced culture, one hour post-induction, four hours post-induction, and overnight were taken to check for protein expression. These samples were run on an SDS-PAGE gel (Figure 3.3). The cultures were harvested by centrifugation at 8,000 rpm for 10 minutes. Cell pellets were stored at -20℃ or resuspended immediately in lysis buffer, at 4℃.

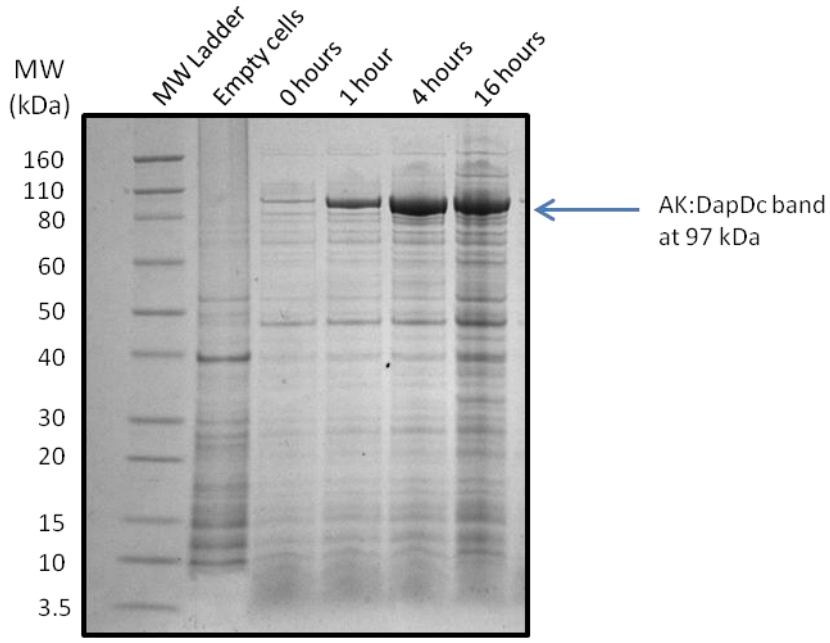


Figure 3.3 - SDS-PAGE of AK:DapDc over-expression.

3.4 Purification

Purification of AK:DapDc from *E. coli* BL21 DE3 cells was achieved after many purification trials, investigating different buffer systems as well as methods of purification. Initial AK:DapDc solubility trials were conducted using PBS pH 7.0. Detergents of varying concentrations (0.1–1% Triton-X 100 and 0.1–1% CHAPS) were added to PBS buffer to aid in solubility. Concentrations of 0.5–1% Triton-X 100 appeared to increase the solubility of His-tagged AK:DapDc. The presence of NaCl or $MgCl_2$ in a lysis buffer was trialled. NaCl gave an increase in solubility while $MgCl_2$ showed no improvement. The addition of PLP or lysine showed no increase in AK:DapDc solubility.

His-Trap purification was first attempted in PBS buffer pH 7.0 on a gravity flow column using a hybrid batch/gravity flow method. AK:DapDc bound to the beads, but would not

elute with 200 nM or 500 mM imidazole, or thrombin cleavage of the His-Tag (Figure 3.4). The presence of Triton-X 100 in the buffers did not change this.

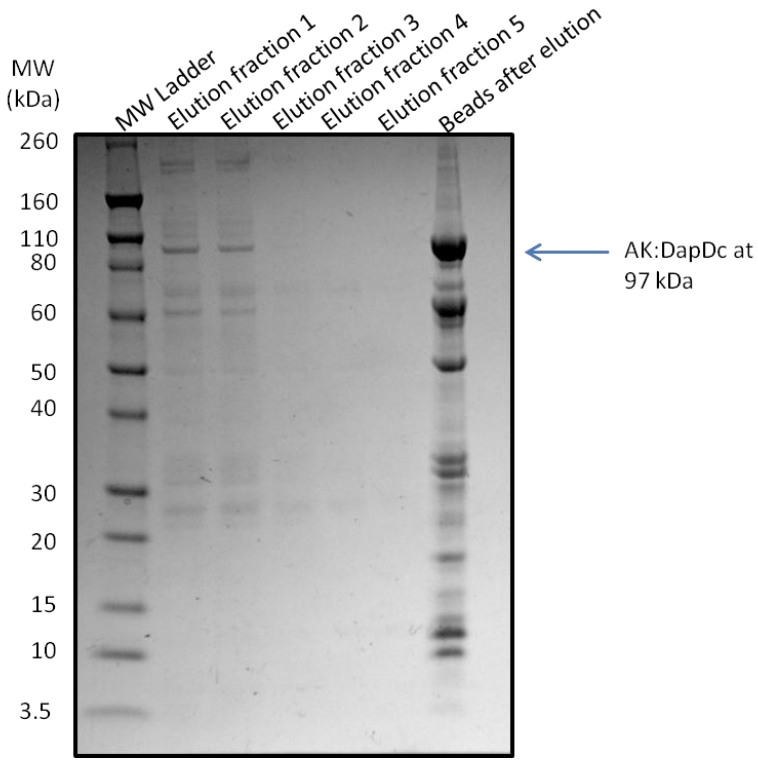


Figure 3.4 - SDS-PAGE analysis of elution from His-Trap column after cleavage of the His-tag by thrombin. The bead fraction was taken after thrombin incubation overnight and subsequent elution.

Anion exchange chromatography was trialled with both His-tagged and untagged AK:DapDc. Most of the AK:DapDc protein did not bind to the column and was present in the flowthrough, using PBS buffer. Small amounts of AK:DapDc bound to the column at pH 8.5 using Tris buffer. The flowthrough was pooled, concentrated, and run through a size exclusion column. The sample did not separate and was very contaminated with other proteins (Figure 3.5). Different buffering conditions were needed to improve the quality of purified AK:DapDc samples.

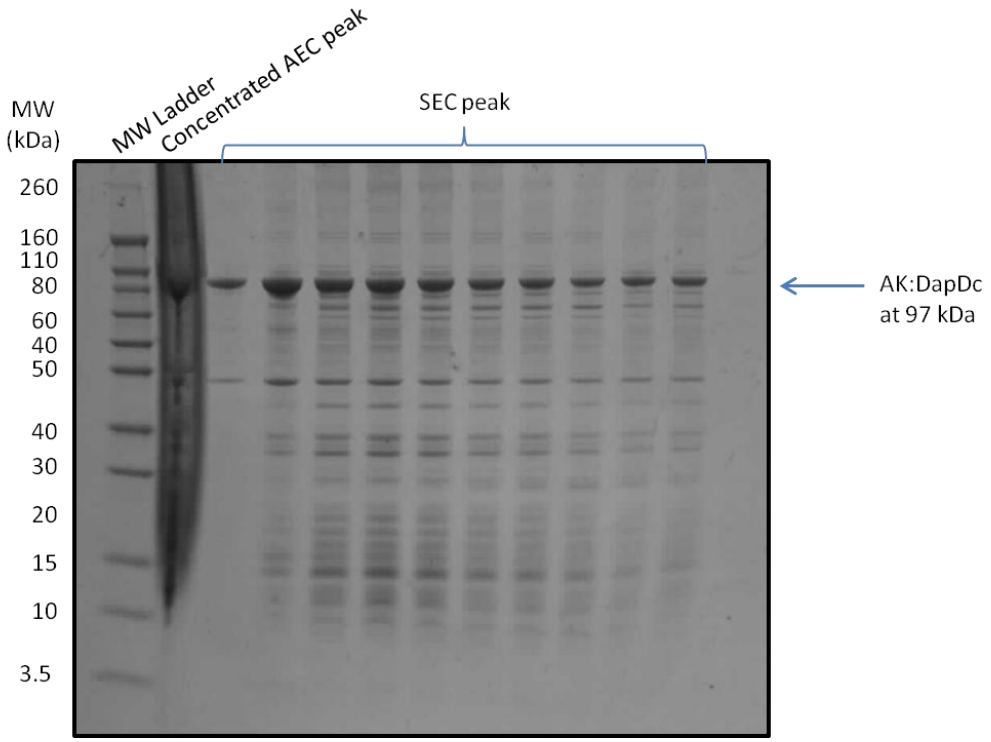


Figure 3.5 - SDS-PAGE analysis of size exclusion. There is no separation of the proteins by size, indicating aggregation.

Fractions of the AEC flowthrough and the size exclusion (SEC) were concentrated and analysed by SDS-PAGE. Both samples were hard spun to remove aggregates. Both spun and un-spun samples were run with and without reducing agent to determine the presence of aggregates or soluble aggregates of AK:DapDc (Figure 3.6). This gel highlights the difference in oligomeric state of AK:DapDc in reducing and non-reducing conditions. The reduced samples show a thicker band of AK:DapDc at 97 kDa, whereas the non-reduced samples have a thinner band at 97 kDa and a number of bands at 200 kDa and higher. These larger proteins are predicted to be AK:DapDc interacting to form higher order structures and aggregates. These results indicated that adding a reducing agent to the purification buffers would aid in preventing aggregation of AK:DapDc.

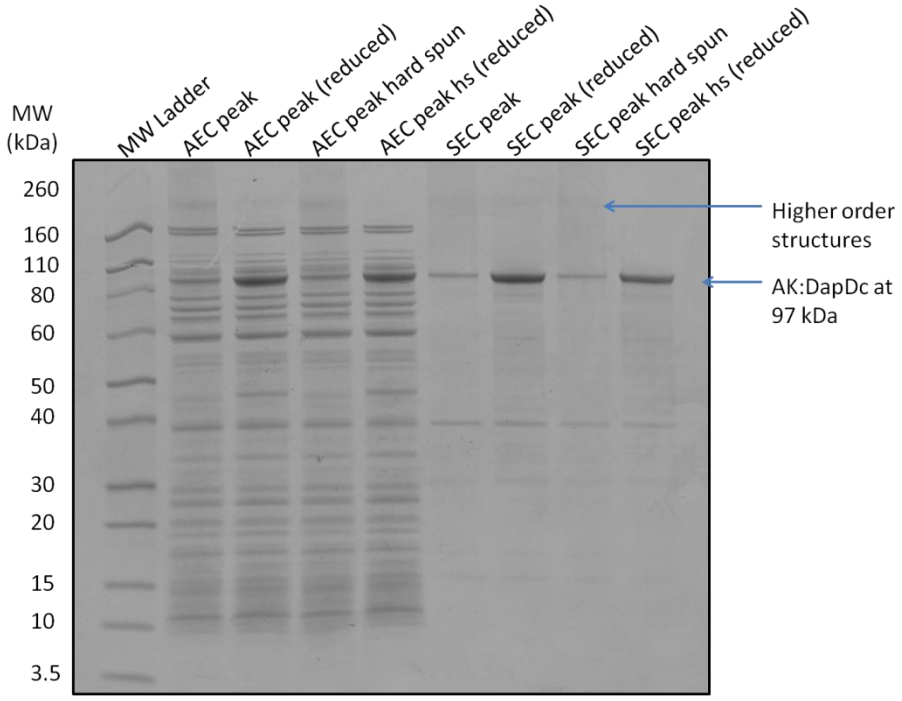


Figure 3.6 - SDS-PAGE analysis of anion exchange and size exclusion chromatography peaks, showing the presence of higher order structures (aggregates) in the non-reduced samples. hs = hard spun.

Differential scanning fluorimetry (DSF) was used to screen buffers and salt concentrations (Table 3.1) according to Seabrook and Newman (2013)[35] with some modifications (Appendix I). The effect of six different buffers in the pH range 4.0-9.0 indicated that a pH less than 6.0 was denaturing, pH 6.0 was not denaturing but gave a much lower melting temperature than pH 7.0-8.0The buffers which gave the highest melting temperatures were citrate pH 6 500 mM NaCl, sodium phosphate pH 7 500 mM NaCl, and Tris pH 7 500mM NaCl. A high salt concentration of 500 mM aided in stability.

Table 3.1 - DSF melting temperatures of AK:DapDc in various buffer systems. The buffers that confer the greatest stability are highlighted in red.

Buffer	pH	Peak 1	Peak 2
Sodium acetate	4	-	-
Sodium acetate	5	38	52
MES	6	48	59/61
Citrate	4	-	-
Citrate	5	39/40	50/51
Citrate	6	52	62
Sodium phosphate	6	51	62
Sodium phosphate	7	52/53	61
Bis-Tris propane	7	51	61
Bis-Tris propane	8	51	60/61
Bis-Tris propane	9	49	61
Tris	7	51/52	64
Tris	8	50/52	63

Taking into account the buffers that DSF indicated to be stabilising, and the effect of a reducing agent on aggregation (Figure 3.6), a buffer system consisting of 20 mM Tris pH 8.5, 200 mM NaCl, 1 mM dithiothreitol, and 5 µM PLP was trialled for use in His-Trap chromatography. This resulted in a successful His-Trap purification of AK:DapDc, which bound to the column and eluted at 43%B (215 mM imidazole). Samples of the soluble

fraction, flowthrough, wash, and peaks were run on a gel against the Novex Pre-stained Protein Standard (Figure 3.7).

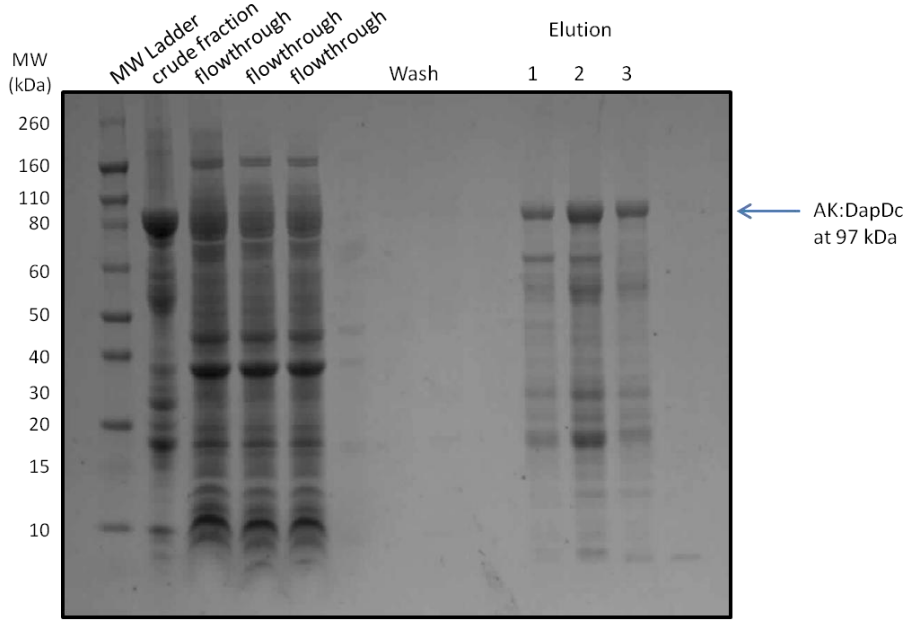


Figure 3.7 - SDS-PAGE analysis of His-Trap chromatography of AK:DapDc.

The His-Trap elution fractions that contained AK:DapDc were concentrated, using a Vivaspin (Sartorius) 6 mL spin concentrator with a 30 kDa cut off, to approximately 2 mL. The sample was loaded onto a 120 mL HiLoad 16/200 Superdex 200 pg (GE Life Sciences) column that was pre-equilibrated with 400 mL of SEC buffer (20 mM Tris pH 7.5, 200 mM NaCl, 5 µM PLP, 1 mM DTT). The sample was loaded onto the column at 0.5 mL/minute. SEC buffer (120 mL) was then run through the column at 0.5 mL/minute. Fractions of 1mL were collected. The protein eluted in multiple peaks (Figure 3.9). Samples of the concentrated protein from His-Trap and the peaks from size exclusion were run on a gel (Figure 3.8).

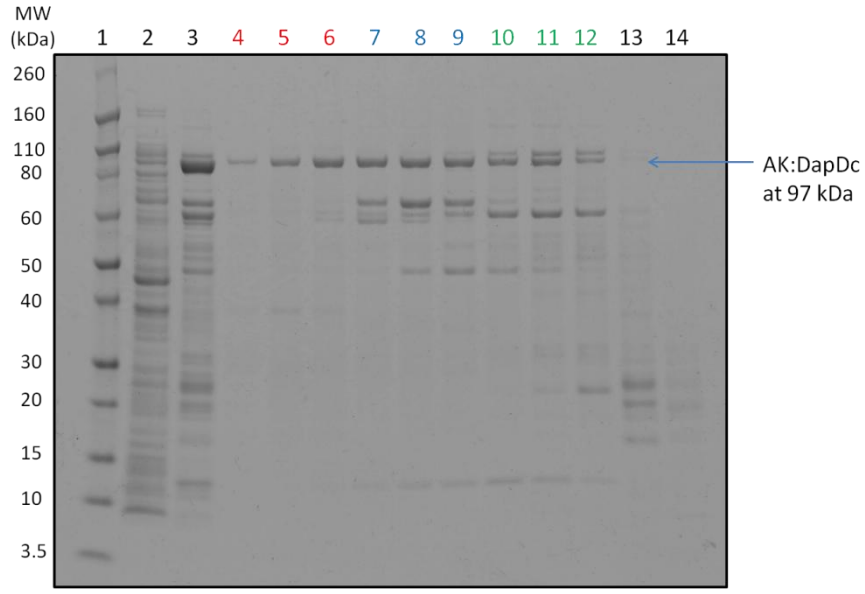


Figure 3.8 - SDS-PAGE analysis of size exclusion. Lanes 4-6: Peak one of size exclusion. Lanes 7-9: Peak two of size exclusion. Lanes 10-12: Peak three of size exclusion.

Lane three of Figure 3.8 shows a sample of the concentrated His-Trap elution peak. Lanes 4-6 are fractions of the first size exclusion elution peak (at approximately 40-50 mL, Figure 3.9), this peak is the most pure but because the protein eluted within the void volume of the column it is predicted to contain aggregates. Lanes 7-9 are fractions from the second elution peak (from approximately 50-60 mL). Lanes 10-12 are fractions from the third elution peak (approximately 65-75 mL). Peaks two and three were used in further characterisation techniques.

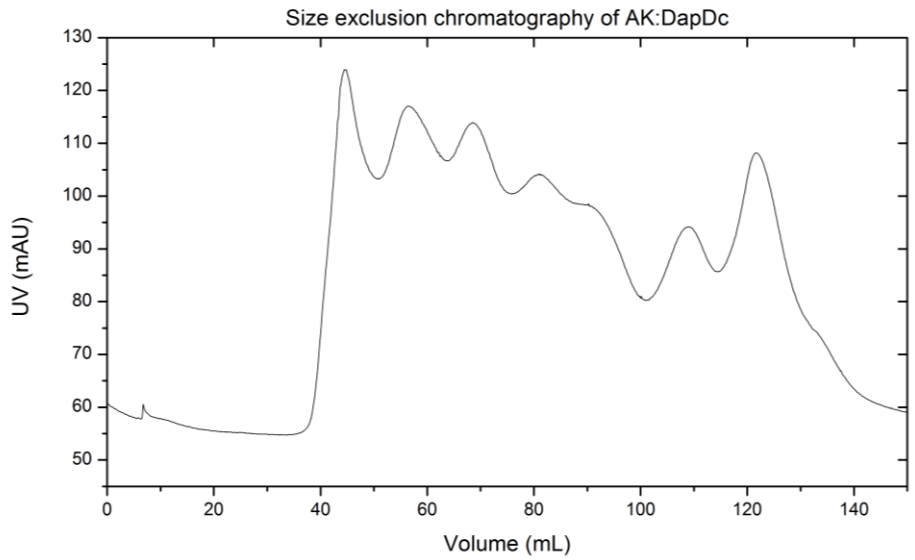


Figure 3.9. Chromatogram of the size exclusion chromatography of AK:DapDc.

The second and third elution peaks from SEC were concentrated using a Vivaspin (Sartorius) 6 mL spin concentrator with a 30 kDa cut off. The protein concentration was measured using a spectrophotometer at 280 nm, with an additional reading at 260 nm to check for DNA contamination. Due to the volumes at which the three peaks of protein eluted, it was suggested that the first peak contained aggregates, the second peak contained large proteins including dimeric AK:DapDc , and the third peak contained smaller proteins including monomeric AK:DapDc.

3.5 Summary of Cloning Expression and Purification

In this study, the XFLM_07435 gene, encoding the AK:DapDc enzyme, was successfully cloned into the pET30ΔSE expression vector. AK:DapDc was then successfully over-expressed in *Escherichia coli* BL21 DE3 cells, after numerous trials. Many more trials were carried out to optimise the purification of AK:DapDc by His-Trap and size exclusion chromatography. It was found that the addition of sodium chloride and the reducing agent

dithiothreitol increased the stability and monodispersity of AK:DapDc. The resulting purified protein is not highly pure, but is a significant improvement from anion exchange chromatography.

References

1. Bi, J. L., Dumenyo, C. K., Hernandez-Martinez, R., Cooksey, D. A., and Toscano, N. C. (2007) Effect of host plant xylem fluid on growth, aggregation, and attachment of Xylella fastidiosa, *Journal of Chemical Ecology 33*, 493-500.
2. Davis, M. J., Purcell, A. H., and Thomson, S. V. (1978) Pierce's Disease of grapevines - Isolation of causal bacterium, *Science 199*, 75-77.
3. Hill, B. L., and Purcell, A. H. (1995) Multiplication and movement of Xylella-fastidiosa within grapevine and 4 other plants, *Phytopathology 85*, 1368-1372.
4. Hopkins, D. L. (1989) Xylella-fastidiosa - Xylem-limited bacterial pathogen of plants, *Annual Review of Phytopathology 27*, 271-290.
5. Purcell, A. H., and Hopkins, D. L. (1996) Fastidious xylem-limited bacterial plant pathogens, *Annual Review of Phytopathology 34*, 131-151.
6. Baumgartner, K., and Warren, J. G. (2005) Persistence of Xylella fastidiosa in riparian hosts near northern California vineyards, *Plant Disease 89*, 1097-1102.
7. Costa, H. S., Raetz, E., Pinckard, T. R., Gispert, C., Hernandez-Martinez, R., Dumenyo, C. K., and Cooksey, D. A. (2004) Plant hosts of Xylella fastidiosa in and near southern California vineyards, *Plant Disease 88*, 1255-1261.
8. Hopkins, D. L., and Adlerz, W. C. (1988) Natural hosts of Xylella-fastidiosa in Florida, *Plant Disease 72*, 429-431.
9. McGaha, L. A., Jackson, B., Bextine, B., McCullough, D., and Morano, L. (2007) Potential Plant Reservoirs for Xylella fastidiosa in South Texas, *American Journal of Enology and Viticulture 58*, 398-401.
10. Mizell, R. F., and French, W. J. (1987) Leafhopper vectors of phony peach disease - Feeding site preference and survival on infected and uninfected peach, and seasonal response to selected host plants, *Journal of Entomological Science 22*, 11-22.
11. Hill, B. L., and Purcell, A. H. (1995) Acquisition and retention of Xylella-fastidiosa by an efficient vector, Graphocephala-atropunctata, *Phytopathology 85*, 209-212.
12. Facincani, A. P., Ferro, J. A., Pizauro, J. M., Pereira, H. A., Lemos, E. G. D., do Prado, A. L., and Ferro, M. I. T. (2003) Carbohydrate metabolism of Xylella fastidiosa: Detection of glycolytic and pentose phosphate pathway enzymes and cloning and expression of the enolase gene, *Genetics and Molecular Biology 26*, 203-211.
13. Simpson, A. J. G., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M. C., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bove, J. M., Briones, M. R. S., Bueno, M. R. P., Camargo, A. A., Camargo, L. E. A., Carraro, D. M., Carrer, H., Colauto, N. B., Colombo, C., Costa, F. F., Costa, M. C. R., Costa-Neto, C. M., Coutinho, L. L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A. P., Ferreira, A. J. S., Ferreira, V. C. A., Ferro, J. A., Fraga, J. S., Franca, S. C., Franco, M. C., Frohme, M., Furlan, L. R., Garnier, M., Goldman, G. H., Goldman, M. H. S., Gomes, S. L., Gruber, A., Ho, P. L., Hoheisel, J. D., Junqueira, M. L., Kemper, E. L., Kitajima, J. P., Krieger, J. E., Kuramae, E. E., Laigret, F., Lambais, M. R., Leite, L. C. C., Lemos, E. G. M., Lemos, M. V. F., Lopes, S. A., Lopes, C. R., Machado, J. A., Machado, M. A., Madeira, A., Madeira, H. M. F., Marino, C. L., Marques, M. V., Martins, E. A. L., Martins, E. M. F., Matsukuma, A. Y., Menck, C. F. M., Miracca, E. C., Miyaki, C. Y., Monteiro-Vitorello, C. B., Moon, D. H., Nagai, M.

A., Nascimento, A., Netto, L. E. S., Nhani, A., Nobrega, F. G., Nunes, L. R., Oliveira, M. A., de Oliveira, M. C., de Oliveira, R. C., Palmieri, D. A., Paris, A., Peixoto, B. R., Pereira, G. A. G., Pereira, H. A., Pesquero, J. B., Quaggio, R. B., Roberto, P. G., Rodrigues, V., Rosa, A. J. D., de Rosa, V. E., de Sa, R. G., Santelli, R. V., Sawasaki, H. E., da Silva, A. C. R., da Silva, A. M., da Silva, F. R., Silva, W. A., da Silveira, J. F., Silvestri, M. L. Z., Siqueira, W. J., de Souza, A. A., de Souza, A. P., Terenzi, M. F., Truffi, D., Tsai, S. M., Tsuhako, M. H., Vallada, H., Van Sluys, M. A., Verjovski-Almeida, S., Vettore, A. L., Zago, M. A., Zatz, M., Meidanis, J., Setubal, J. C., and Xylella fastidiosa Consortium, O. (2000) The genome sequence of the plant pathogen Xylella fastidiosa, *Nature 406*, 151-157.

14. Lemos, E. G. d. M., Alves, L. M. C., and Campanharo, J. C. (2003) Genomics-based design of defined growth media for the plant pathogen Xylella fastidiosa, *FEMS Microbiology Letters 219*, 39-45.
15. Fondi, M., Brilli, M., and Fani, R. (2007) On the origin and evolution of biosynthetic pathways: integrating microarray data with structure and organization of the common pathway genes, *Bmc Bioinformatics 8*.
16. Kotaka, M., Ren, J., Lockyer, M., Hawkins, A. R., and Stammers, D. K. (2006) Structures of R- and T-state Escherichia coli aspartokinase III - Mechanisms of the allosteric transition and inhibition by lysine, *Journal of Biological Chemistry 281*, 31544-31552.
17. Wampler, D. E., and Westhead, E. W. (1968) 2 Aspartokinases from Escherichia coli nature of inhibition and molecular changes accompanying reversible inactivation, *Biochemistry 7*, 1661-&.
18. Black, S., and Wright, N. G. (1955) Beta-aspartokinase and beta-aspartyl phosphate, *Journal of Biological Chemistry 213*, 27-38.
19. Tunca, S., Yilmaz, E. I., Piret, J., Liras, P., and Ozcengiz, G. (2004) Cloning, characterization and heterologous expression of the aspartokinase and aspartate semialdehyde dehydrogenase genes of cephamycin C-producer Streptomyces clavuligerus, *Research in Microbiology 155*, 525-534.
20. Schuldt, L., Suchowersky, R., Veith, K., Mueller-Dieckmann, J., and Weiss, M. S. (2011) Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of the regulatory domain of aspartokinase (Rv3709c) from Mycobacterium tuberculosis, *Acta Crystallographica Section F-Structural Biology and Crystallization Communications 67*, 380-385.
21. Zhang, W. W., Jiang, W. H., Zhao, G. P., Yang, Y. L., and Chiao, J. S. (2000) Expression in Escherichia coli, purification and kinetic analysis of the aspartokinase and aspartate semialdehyde dehydrogenase from the rifamycin SV-producing Amycolatopsis mediterranei U32, *Applied Microbiology and Biotechnology 54*, 52-58.
22. Rosner, A. (1975) Control of Lysine biosynthesis in Bacillus-subtilis - Inhibition of Diaminopimelate decarboxylase by Lysine, *Journal of Bacteriology 121*, 20-28.
23. Robin, A. Y., Cobessi, D., Curien, G., Robert-Genthon, M., Ferrer, J. L., and Dumas, R. (2010) A New Mode of Dimerization of Allosteric Enzymes with ACT Domains Revealed by the Crystal Structure of the Aspartate Kinase from Cyanobacteria, *Journal of Molecular Biology 399*, 283-293.
24. Falcoz-Kelly, F., van Rapenbusch, R., and Cohen, G. N. (1969) The methionine-repressible homoserine dehydrogenase and aspartokinase activities of Escherichia coli K 12. Preparation of the homogeneous protein catalyzing the two activities.

Molecular weight of the native enzyme and of its subunits, *European Journal of Biochemistry 8*, 146-152.
25. Truffa-Bachi, P., van Rapenbusch, R., Janin, J., Gros, C., and Cohen, G. N. (1968) The Threonine-Sensitive Homoserine Dehydrogenase and Aspartokinase Activities of Escherichia coli K 12, *European Journal of Biochemistry 5*, 73-80.
26. White, P. J., and Kelly, B. (1965) Purification and properties of Diaminopimelate decarboxylase from Escherichia coli, *Biochemical Journal 96*, 75-&.
27. Fogle, E. J., and Toney, M. D. (2011) Analysis of catalytic determinants of diaminopimelate and ornithine decarboxylases using alternate substrates, *Biochimica Et Biophysica Acta-Proteins and Proteomics 1814*, 1113-1119.
28. Lakshman, M., Shenoy, B. C., and Rao, M. R. R. (1981) Purification and properties of Diaminopimelate decarboxylase of Micrococcus-glutamicus, *Journal of Biosciences 3*, 89-103.
29. Cheraghi, S., Akbarzade, A., Farhangi, A., Chiani, M., Saffari, Z., Ghassemi, S., Rastegari, H., and Mehrabi, M. R. (2010) Improved production of l-lysine by over-expression of meso-diaminopimelate decarboxylase enzyme of Corynebacterium glutamicum in Escherichia coli, *Pakistan Journal of Biological Sciences 13*, 504-508.
30. White, P. J., Suffling, A., Kelly, B., and Work, E. (1964) Variation of activity of bacterial Diaminopimelate decarboxylase under different conditions of growth, *Biochemical Journal 91*, 600-&.
31. Bourot, S., Sire, O., Trautwetter, A., Touzé, T., Wu, L. F., Blanco, C., and Bernard, T. (2000) Glycine betaine-assisted protein folding in a lysA mutant of Escherichia coli, *Journal of Biological Chemistry 275*, 1050-1056.
32. Hu, T. C., Wu, D. L., Chen, J., Ding, J. P., Jiang, H. L., and Shen, X. (2008) The catalytic intermediate stabilized by a "Down" active site loop for diaminopimelate decarboxylase from Helicobacter pylori - Enzymatic characterization with crystal structure analysis, *Journal of Biological Chemistry 283*, 21284-21293.
33. Weyand, S., Kefala, G., Svergun, D. I., and Weiss, M. S. (2009) The three-dimensional structure of diaminopimelate decarboxylase from Mycobacterium tuberculosis reveals a tetrameric enzyme organisation, *Journal of structural and functional genomics 10*, 209-217.
34. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *Journal of Molecular Biology 215*, 403-410.
35. Seabrook, S. A., and Newman, J. (2013) High-Throughput Thermal Scanning for Protein Stability: Making a Good Technique More Robust, *Acs Combinatorial Science 15*, 387-392.

4. Kinetic Assaying of AK:DapDc Using a Coupled Assay

4.1 Introduction

A major focus of this study is to confirm the predicted functions of Aspartate kinase:Diaminopimelate decarboxylase (AK:DapDc). Whether or not both reactions are carried out by AK:DapDc will indicate its role in the cell. *Xylella fastidiosa* has two genes encoding aspartate kinase, both of which encode bifunctional enzymes[1]. If both of these are functional, this suggests that they have different cellular roles and may be regulated by different pathway end-products. *X. fastidiosa* has only one gene encoding diaminopimelate decarboxylase[1]. DapDc is inhibited by lysine in bacteria, so AK:DapDc is likely to also be inhibited by lysine [2]. If both functions catalyse their predicted reactions and both are inhibited by lysine, then this confirms the hypothesised regulatory role of AK:DapDc in lysine biosynthesis.

There are three common methods to assay for either aspartate kinase or diaminopimelate decarboxylase activity, with the preferred method being coupled enzyme assays in each case. There are also colorimetric, spectrophotometric, or radiolabelling endpoint assays. The different types of assays measure different aspects of a reaction, so choosing the right assay method is important. For this study, measuring the initial rate of reaction is important, because kinetic parameters such as maximum velocity (V_{max}), Michaelis-Menten constant (K_M), and catalytic efficiency (kcat) can be calculated from the initial rate. These parameters contribute to the characterisation of the AK:DapDc enzyme.

4.1.1 Aspartate Kinase Assay

Aspartate kinase catalyses the phosphorylation of aspartate, using ATP to produce aspartyl phosphate and ADP (Figure 1.5). Aspartate kinase activity is usually assayed using the spectrophotometric aspartyl-hydroxamate endpoint assay[3], the aspartic semialdehyde dehydrogenase coupled enzyme assay[4], or the pyruvate kinase coupled enzyme assay[4]. The aspartyl-hydroxamate assay uses hydroxylamine to measure the formation of a hydroxamic acid spectrophotometrically at a wavelength of 540 nm[5]. Hydroxamic acid is formed when the hydroxylamine reacts with aspartyl phosphate, the product of the aspartate kinase reaction. This is a discontinuous spectrophotometric assay that measures product production at different time points after the reaction is stopped with ferric chloride[3]. This method does not give an initial rate of reaction, so it is not ideal for measuring the activity of AK:DapDc and was not used in this study.

The aspartic-semialdehyde dehydrogenase (ASADH) coupled assay couples the production of aspartyl phosphate by aspartate kinase, with the conversion of aspartyl phosphate to aspartate semialdehyde by aspartate semialdehyde dehydrogenase[6]. The formation of aspartate semialdehyde occurs with concomitant oxidation of NADPH to $NADP^{+}$[4]. This oxidation can be monitored spectrophotometrically at a wavelength of 340 nm. In contrast to the aspartyl-hydroxamate assay, the ASADH assay is continuous. Thus an initial rate of reaction can be accurately calculated.

The pyruvate kinase assay couples two other enzymatic reactions to the aspartate kinase reaction (Figure 4.1)[7]. When aspartate is converted to aspartyl phosphate, ATP is used as the phosphate source. The ADP produced from the reaction is then used by pyruvate kinase to convert phosphoenol pyruvate into pyruvate and ATP. Pyruvate is then converted to lactate by lactate dehydrogenase with concomitant oxidation of NADH to NAD^{+}. The

reduction of NADH is monitored spectrophotometrically at 340 nm. This is another spectrophotometric continuous assay which facilitates the accurate measurement of initial rates.

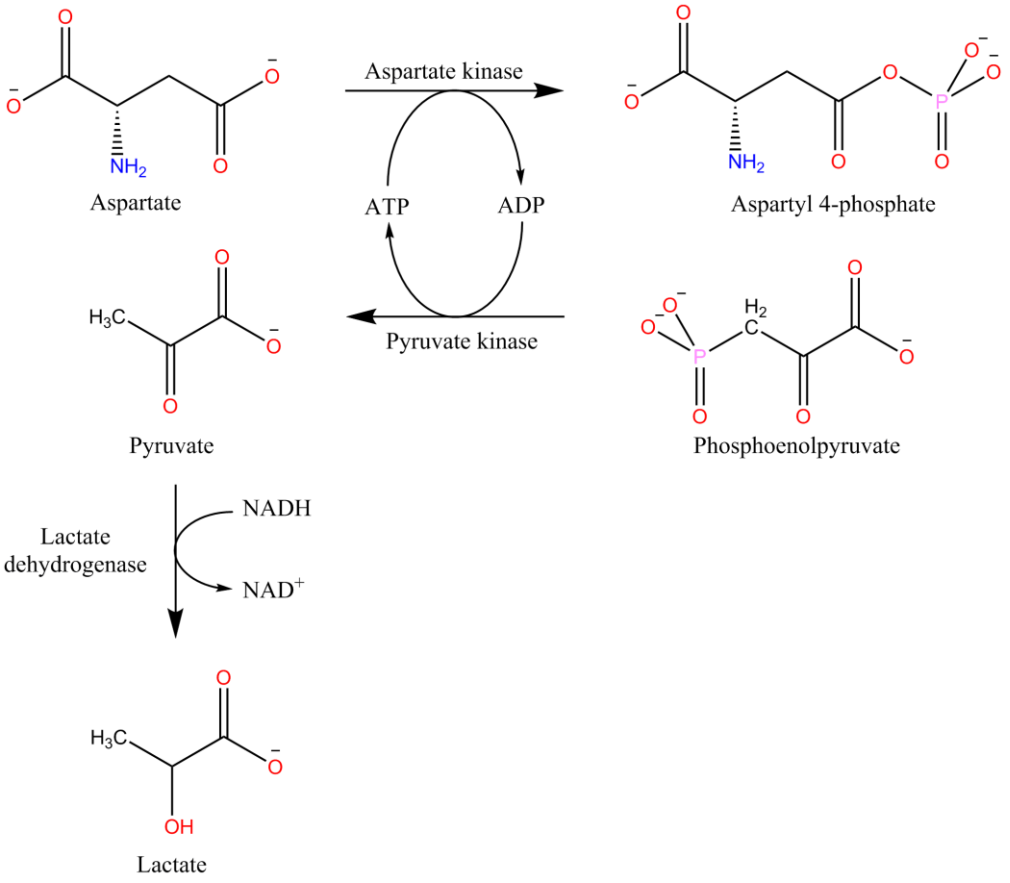


Figure 4.1 - Pyruvate kinase assay for aspartase kinase activity. This assay couples the AK reaction to pyruvate kinase and lactate dehydrogenase. The disappearance of NADH is measured at 340 nm. Drawn in Chemdraw.

For this study, the pyruvate kinase coupled assay was determined to be the most ideal. However, due to time constraints and availability of certain substrates, the activity of aspartate kinase was not measured.

4.1.2 Diaminopimelate Decarboxylase Assay

DapDc catalyses the decarboxylation of *meso*-diaminopimelate to give L-lysine and carbon dioxide. DapDc activity is commonly assayed by a radiochemical (C^{14}) assay, a ninhydrin-based colorimetric assay, or coupled enzyme assays. The radiochemical assay uses radio-labelled C^{14} diaminopimelate and measures the evolution of C^{14}-containing carbon dioxide directly[8].

The ninhydrin-based method is a discontinuous colorimetric assay[9]. In this assay, samples of the reaction are taken at different time points and the reaction is stopped with glacial acetic acid. Ninhydrin reagent is added and incubated at 37°C for 90 minutes. The ninhydrin reagent forms coloured products with short chain amino acids containing two amino groups[10]. It turns purple with diaminopimelate and yellow with lysine. The absorbance of the coloured sample is measured at 440 nm and correlates with concentration of lysine present. This method is time consuming and does not give initial rate data, so it is not ideal for estimation of AK:DapDc activity. Therefore, this assay was not used in this study.

The coupled enzyme assays link additional reactions to the conversion of *meso*-diaminopimelate to lysine and carbon dioxide. In the common two-enzyme assay, the carbon dioxide produced from the decarboxylation of *meso*-diaminopimelate is used by PEP carboxylase to convert phosphoenolpyruvate to oxaloacetate. Malate dehydrogenase then converts oxaloacetate and NADH to malate and NAD^+. The disappearance of NADH is monitored at 340 nm.

The coupled assay used in this study couples the DapDc reaction to the reaction of saccharopine dehydrogenase (SDH) (Figure 4.2)[11]. The DapDc reaction produces lysine, which reacts with α-ketoglutarate, catalysed by saccharopine dehydrogenase, to give

saccharopine. This reaction occurs with concomitant oxidation of NADH to NAD^+. The oxidation of NADH is monitored continuously at 340 nm, so this assay also gives an accurate calculation of the initial rate of reaction. This assay also requires only one coupling enzyme and one set of substrates.

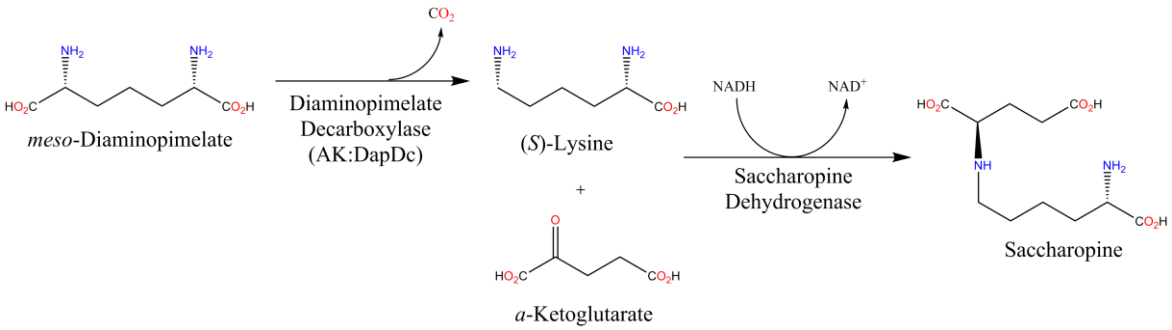


Figure 4.2 - Saccharopine dehydrogenase activity assay reaction

4.2 Over-expression of Saccharopine Dehydrogenase from *Saccharomyces cerevisiae*

Escherichia coli BL21 cells were transformed with pET16b-SDH vector (containing the gene for SDH from *Saccharomyces cerevisiae*) according to Section 2.2.7. The transformation resulted in many colonies. Several colonies were pre-cultured and SDH was expressed according to Sections 2.3.1-2.3.3. A cell sample from 16 hours post-induction was analysed via SDS-PAGE and this confirmed that SDH over-expressed (Figure 4.3). Optimal conditions for over-expression of SDH included growth of the cells to an OD600 of 0.7-0.9 at 37°C, induction with 1 mM IPTG and overnight incubation at 26°C.

4.3 Purification of SDH from *E. coli* BL21 Cells

SDH was purified according to Xu *et al.* (2006) with some modifications according to advice from Matthew Perugini's lab (unpublished data)[12]. Figure 4.3 shows the SDS-PAGE analysis of the purification. The total fraction is a sample of *E. coli* BL21 cells taken before

harvesting, which indicates over-expression of SDH, at 44 kDa. Lysed and centrifuged SDH bound well to the column and very little eluted during the washes. Wash 1 is a fraction from the collected flowthrough of the first wash. Wash 2 is a fraction of the elution due to a low imidazole (20 mM) wash. SDH eluted with 150 mM imidazole[12], as is shown by Elution fractions 2-9 in Figure 4.3. Approximately 15 elution fractions were collected, with most of the SDH eluting in the first 10 fractions. There is still a considerable amount of SDH bound to the beads, as the Beads fraction was taken after elution of SDH.

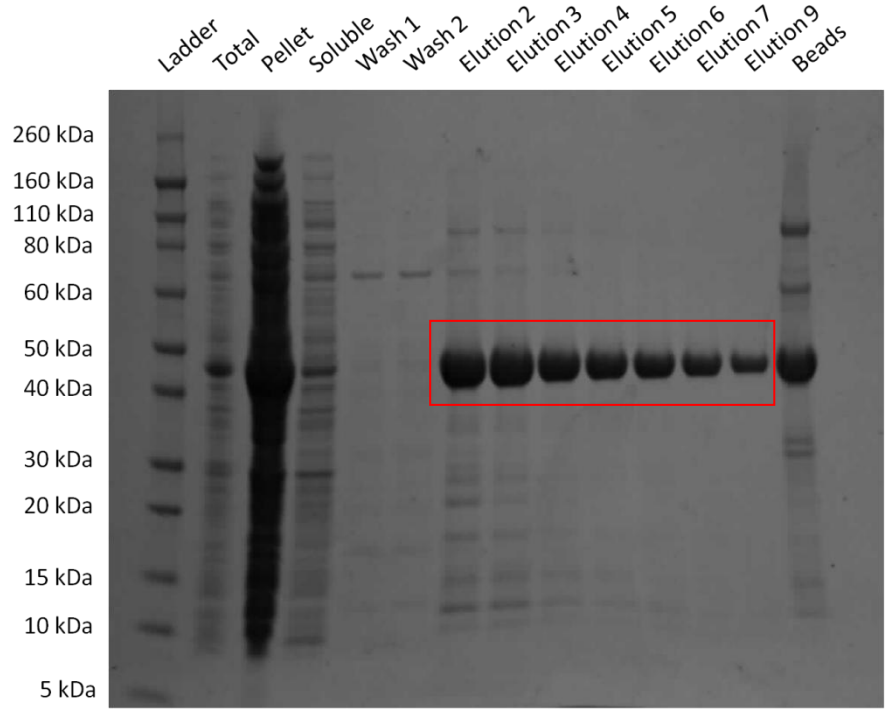


Figure 4.3 - SDS-PAGE gel of SDH purification. SDH elution fractions are as indicated by the red box.

4.4 Kinetic Assay

The SDH assay used in this project is similar to that of Xu *et al.* (2006) and Ogawa & Fujioka (1978), but uses a different buffer system and temperature (200 mM Tris buffer pH 8.0 at 37°C) following recommendations from Matthew Perugini's lab (unpublished

data)[12, 13]. Xu *et al.* (2006) used 100 mM HEPES pH 7.0 at 25℃ and Ogawa and Fujioka used 100 mM potassium phosphate pH 6.8 at 22°C.

4.4.1 [SDH] vs. Rate of Reaction

First, the change in initial rate of conversion of lysine and α-ketoglutarate to saccharopine with increasing SDH concentration was measured. The purpose of this is to determine the concentration of SDH which is proportional to the rate of reaction. No measurements were taken above 1.5 μM SDH as the initial rate became too difficult to measure and calculate accurately. Figure 3 shows that at concentrations of SDH higher than 1.5 μM, the reaction rate appears to level off.

The SDH concentration proportional to the rate of reaction is 0.25 μM (Figure 4.4). This was determined using Figure 4.4. This SDH concentration was used to further determine the kinetic parameters of SDH.

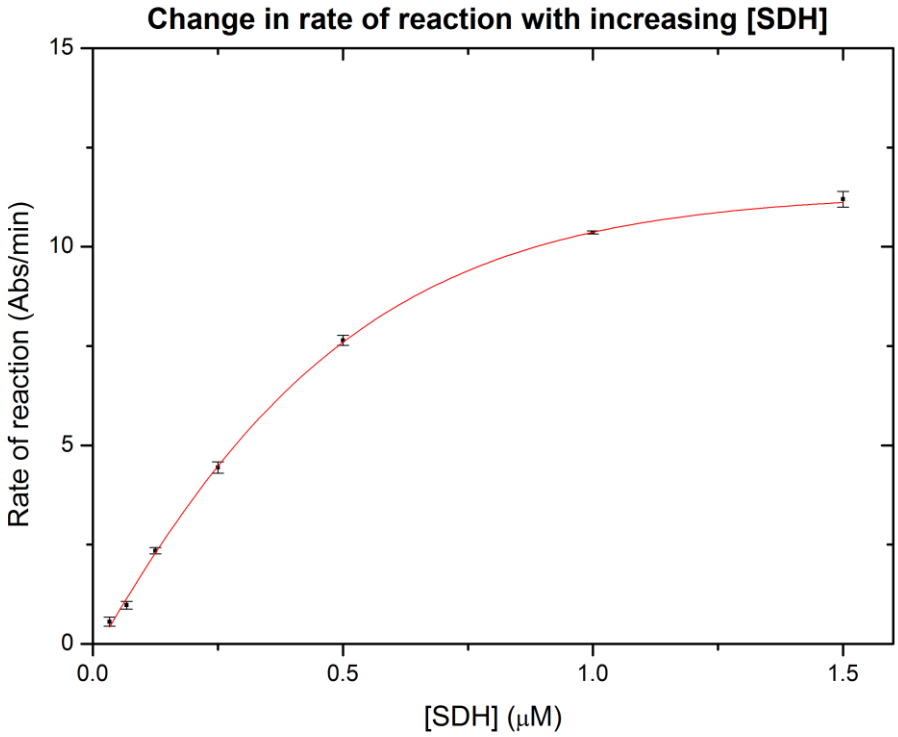


Figure 4.4 - Graph of the increase in rate of reaction with increase in [SDH]. Each data point is an average of three replicates. The error bars represent standard error of the mean. Analysed in sedfit.

4.4.2 [Lysine] vs. Rate of Reaction

The next set of assays estimated the increase in rate of reaction with increasing lysine concentration, while keeping the SDH concentration constant. The purpose of this step was to determine $K_{M(Lys)}$ and V_{max} of SDH purified in Section 4.3. The concentration of lysine ranged from 1-100 mM (1 mM, 2 mM, 5 mM, 10 mM, 15 mM, 25 mM, 50 mM, 100 mM) while keeping the α-ketoglutarate (α-KG) concentration at 25 mM. This concentration of α-ketoglutarate was used because it is at least ten times the $K_{M(α-KG)}$ reported for *Sc*SDH and other SDH enzymes. Xu *et al* (2006) reports the $K_{M(α-KG)}$ to be 0.11±0.03 mM, while Zabriskie & Jackson (2000) report a $K_{M(α-KG)}$ of 0.55 mM[12, 14]. The SDH concentration was

0.25 μM, as this was estimated to be in the range of concentrations where the rate of reaction is proportional to the concentration of SDH present (See previous section).

The lysine titration assay (Figure 4.5) shows that the rate of reaction approaches the maximum velocity after 50 mM lysine. In the activity assay, lysine is produced by AK:DapDc and needs to be immediately converted by SDH to saccharopine, so the affinity of SDH for lysine is important to estimate. The affinity for lysine ($K_{M(Lys)}$) of the SDH expressed and purified in this study is 19.3±2.3 mM. The maximum velocity of this enzyme is 7.8±0.3 abs/min.

The SDH purified and used in this study has a lower affinity ($K_{M(Lys)}$) for lysine than that of Xu *et al.* (2006)[12]. Xu *et al.* estimated the $K_{M(Lys)}$ to be 1.1 ± 0.2 mM. This difference could be due to using different expression and purification conditions.

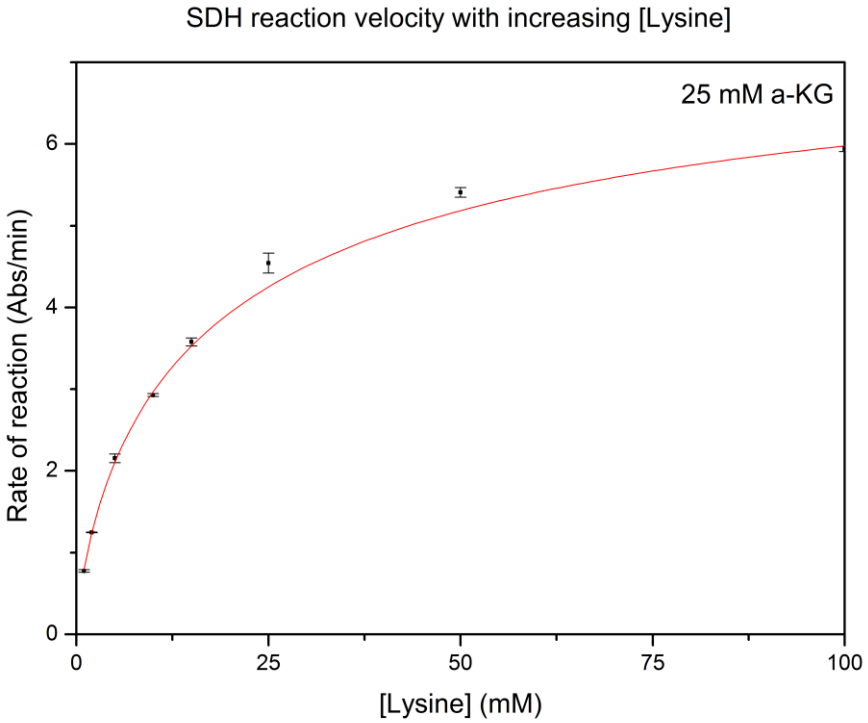


Figure 4.5 - Lysine titration assay. Increase in initial rate of reaction with respect to increased lysine concentration, with 0.25µM SDH. Each data point is an average of three replicates. Error bars represent the standard error of the mean. Analysed in sedfit.

4.4.3 AK:DapDc Activity

To test for AK:DapDc, the assay was set up as per Section 1.3.7.3, but without lysine, and with a final SDH concentration of 3 µM. The assay mix contained all the substrates (excluding lysine), but neither of the enzymes. After incubating each cuvette of assay mix (~950 µL) to 37°C, the SDH was added and the absorbance was monitored for approximately 10 seconds to ensure there was no SDH activity with the other substrates. Then, 20 µL of approximately 1 mg/mL AK:DapDc was added and the absorbance at 340 nm was monitored for 2 minutes.

Several different samples of AK:DapDc were assayed: crude/soluble fraction (containing all soluble cellular proteins), the AK:DapDc-containing pooled elution peak from His-Trap

chromatography, and the pooled concentrated AK:DapDc elution from SEC. Both fresh (purified that day) and frozen (with and without glycerol) samples of purified AK:DapDc were assayed. Activity was only detected in the crude fraction sample, but as that fraction contained all soluble cellular proteins, and presumably also lysine (at a physiological concentration) it cannot be confirmed that AK:DapDc was active.

Thus, AK:DapDc may not be active in 200 mM Tris pH 8.0 at 37°C. *X. fastidiosa*, the bacterium that AK:DapDc originates from, grows optimally at 26-28°C[15]. This may be the optimal temperature for AK:DapDc activity. The melting temperature for AK:DapDc in Tris buffer is much higher than 37°C (See DSF results, section 3.4.1.1), so the protein is not likely to be denatured due to temperature.

4.5 Summary of Kinetics

In this study, the *S. cerevisiae* SDH was successfully over-expressed and purified (Figure 4.3). The $K_{M(Lys)}$ and V_{max} of SDH were calculated and compared to those found by other studies. The difference between the parameters calculated in this study, and those calculated by Xu *et al.* (2006)[12] may have arisen from differences in the expression system, purification conditions, or assay conditions. The next step for this assay is to calculate the $K_{M(\alpha\text{-KG})}$ for SDH in order to ensure that the concentration of α-KG included in the assay is in fact sufficient to be in excess. If active AK:DapDc is obtained the concentration of SDH that is in excess compared to AK:DapDc could be determined. This would prevent SDH from being a rate limiting factor in the assay.

References

1. Simpson, A. J. G., Reinach, F. C., Arruda, P., Abreu, F. A., Acencio, M., Alvarenga, R., Alves, L. M. C., Araya, J. E., Baia, G. S., Baptista, C. S., Barros, M. H., Bonaccorsi, E. D., Bordin, S., Bove, J. M., Briones, M. R. S., Bueno, M. R. P., Camargo, A. A., Camargo, L. E. A., Carraro, D. M., Carrer, H., Colauto, N. B., Colombo, C., Costa, F. F., Costa, M. C. R., Costa-Neto, C. M., Coutinho, L. L., Cristofani, M., Dias-Neto, E., Docena, C., El-Dorry, H., Facincani, A. P., Ferreira, A. J. S., Ferreira, V. C. A., Ferro, J. A., Fraga, J. S., Franca, S. C., Franco, M. C., Frohme, M., Furlan, L. R., Garnier, M., Goldman, G. H., Goldman, M. H. S., Gomes, S. L., Gruber, A., Ho, P. L., Hoheisel, J. D., Junqueira, M. L., Kemper, E. L., Kitajima, J. P., Krieger, J. E., Kuramae, E. E., Laigret, F., Lambais, M. R., Leite, L. C. C., Lemos, E. G. M., Lemos, M. V. F., Lopes, S. A., Lopes, C. R., Machado, J. A., Machado, M. A., Madeira, A., Madeira, H. M. F., Marino, C. L., Marques, M. V., Martins, E. A. L., Martins, E. M. F., Matsukuma, A. Y., Menck, C. F. M., Miracca, E. C., Miyaki, C. Y., Monteiro-Vitorello, C. B., Moon, D. H., Nagai, M. A., Nascimento, A., Netto, L. E. S., Nhani, A., Nobrega, F. G., Nunes, L. R., Oliveira, M. A., de Oliveira, M. C., de Oliveira, R. C., Palmieri, D. A., Paris, A., Peixoto, B. R., Pereira, G. A. G., Pereira, H. A., Pesquero, J. B., Quaggio, R. B., Roberto, P. G., Rodrigues, V., Rosa, A. J. D., de Rosa, V. E., de Sa, R. G., Santelli, R. V., Sawasaki, H. E., da Silva, A. C. R., da Silva, A. M., da Silva, F. R., Silva, W. A., da Silveira, J. F., Silvestri, M. L. Z., Siqueira, W. J., de Souza, A. A., de Souza, A. P., Terenzi, M. F., Truffi, D., Tsai, S. M., Tsuhako, M. H., Vallada, H., Van Sluys, M. A., Verjovski-Almeida, S., Vettore, A. L., Zago, M. A., Zatz, M., Meidanis, J., Setubal, J. C., and Xylella fastidiosa Consortium, O. (2000) The genome sequence of the plant pathogen Xylella fastidiosa, *Nature 406*, 151-157.
2. White, P. J., and Kelly, B. (1965) Purification and properties of Diaminopimelate decarboxylase from Escherichia coli, *Biochemical Journal 96*, 75-&.
3. Black, S., and Wright, N. G. (1955) BETA-ASPARTOKINASE AND BETA-ASPARTYL PHOSPHATE, *Journal of Biological Chemistry 213*, 27-38.
4. Wampler, D. E., and Westhead, E. W. (1968) 2 Aspartokinases from Escherichia coli nature of inhibition and molecular changes accompanying reversible inactivation, *Biochemistry 7*, 1661-&.
5. Stadtman, E. R., Cohen, G. N., Robichonszulmajster, H. D., and Lebras, G. (1961) Feed-back inhibition and repression of Aspartokinase activity in Escherichia coli and Saccharomyces cerevisiae, *Journal of Biological Chemistry 236*, 2033-&.
6. Angeles, T. S., Hunsley, J. R., and Viola, R. E. (1992) Reversal of enzyme regiospecificity with alternative substrates for aspartokinase I from Escherichia coli, *Biochemistry 31*, 799-805.
7. Chassagnole, C., Rais, B., Quentin, E., Fell, D. A., and Mazat, J. P. (2001) An integrated study of threonine-pathway enzyme kinetics in Escherichia coli, *Biochemical Journal 356*, 415-423.
8. Scriven, F., Wlasichuk, K. B., and Palcic, M. M. (1988) A CONTINUAL SPECTROPHOTOMETRIC ASSAY FOR AMINO-ACID DECARBOXYLASES, *Analytical Biochemistry 170*, 367-371.
9. Momany, C., Levdikov, V., Blagova, L., and Crews, K. (2002) Crystallization of diaminopimelate decarboxylase from Escherichia coli, a stereospecific D-amino-acid decarboxylase, *Acta Crystallographica Section D-Biological Crystallography 58*, 549-552.

10. Work, E. (1957) REACTION OF NINHYDRIN IN ACID SOLUTION WITH STRAIGHT-CHAIN AMINO ACIDS CONTAINING 2 AMINO GROUPS AND ITS APPLICATION TO THE ESTIMATION OF ALPHA-EPSILON-DIAMINOPIMELIC ACID, *Biochemical Journal 67*, 416-423.
11. Laber, B., and Amrhein, N. (1989) A SPECTROPHOTOMETRIC ASSAY FOR MESO-DIAMINOPIMELATE DECARBOXYLASE AND L-ALPHA-AMINO-EPSILON-CAPROLACTAM HYDROLASE, *Analytical Biochemistry 181*, 297-301.
12. Xu, H. Y., West, A. H., and Cook, P. F. (2006) Overall kinetic mechanism of saccharopine dehydrogenase from Saccharomyces cerevisiae, *Biochemistry 45*, 12156-12166.
13. Ogawa, H., and Fujioka, M. (1978) Purification and characterization of saccharopine dehydrogenase from baker's yeast, *Journal of Biological Chemistry 253*, 3666-3670.
14. Zabriskie, T. M., and Jackson, M. D. (2000) Lysine biosynthesis and metabolism in fungi, *Natural Product Reports 17*, 85-97.
15. Janse, J. D., and Obradovic, A. (2010) Xylella fastidiosa: Its biology, diagnosis, control and risks, *Journal of Plant Pathology 92*, S35-S48.

5. Biophysical Characterization

5.1 Introduction

Protein structure and function are interlinked, therefore structural and biophysical characterisation is key to understanding a protein's cellular role. Biophysical techniques provide information on the structural properties of the protein.

Circular dichroism (CD) evaluates the secondary structures present in the protein and can confirm the proportions of α-helical, β-strand/sheet, and unstructured regions[1]. Web-based programs (e.g. PSIPRED) can be used to computationally predict the secondary structures present[2]. This computational prediction can show the presence of fold patterns, but also show where highly unstructured regions are present. These unstructured regions may affect protein solubility and stability, which need to be taken into consideration during expression and purification. These techniques can be supported by experimental and theoretical techniques that determine or predict tertiary and quaternary structure.

Generating homology models is a technique used to model the three-dimensional (3D) structure of a protein when there is no experimental 3D structure available. The technique is based on sequence identity of a target amino acid sequence to a database of amino acid sequences of proteins with a known 3D structure. The purpose of homology models is to predict the structure as a guide for further applications such as mutagenesis studies and protein engineering, as well as supporting experimental x-ray crystallographic structures[3]. SWISS-MODEL is a fully automated homology modelling server that has been available for 20 years[4]. Recently, it has been improved to be able to model quaternary structure

(oligomeric state) in addition to tertiary structure (3D protein folding)[4]. This server was selected to construct homology models in this study.

Analytical Ultracentrifugation (AUC) is an experimental technique which estimates physical properties of proteins by measuring the effect of a centrifugal force on the target protein[5]. AUC estimates molecular weight, sedimentation coefficient/velocity, oligomeric state and shape of proteins[6]. This technique is also used to determine the effects of DNA or small molecule binding to the target protein[7]. In this study, AUC was used to determine the molecular weight of AK:DapDc and its quaternary structure in solution.

X-ray crystallography is an integral part of characterising a protein. It provides a 3D atomic structure of the protein and can also show oligomeric state. A 3D structure with ligands bound can provide an insight into the enzymatic reaction or inhibition of the protein. In order to obtain high quality diffracting crystals, screening and optimisation of crystallisation conditions must be trialled. During this project four crystal screens were trialled at two temperatures, using protein purified using anion exchange chromatography and SEC. The preferred method of purification (His-Trap followed by SEC) did not yield enough protein for further trials. However, this small scale trial still provides useful information about which crystallisation conditions to explore further.

5.2 Circular Dichroism (CD)

CD was used to determine the secondary structure and confirm the predicted proportions of secondary structures present in AK:DapDc.

First, buffer tests at 20°C determined which buffer was the best to use (See section 2.4.3). The ideal buffer gives a reading of < 600 high tension voltage (HT [V]) for the entire

wavelength range measured. An HT voltage of ≥ 600 V indicates the spectrophotometer sensor is saturated and measurements taken at that wavelength are not reliable[8]. Salts and other buffer additives affect the accuracy of CD measurements at lower wavelengths, by increasing the HT voltage to above 600 V at wavelengths less than 200 nm[9].

However, AK:DapDc was purified in a buffer containing salt (NaCl), a reducing agent (1 mM DTT), and the cofactor (5 μM PLP). Careful optimisation showed that Tris buffer (5 mM pH 7.5) was found to be the most ideal concentration. Individually, 1 mM DTT or 5 μM PLP did not affect the HT voltage significantly, but together they had a noticeable effect. Sodium chloride was not added to the CD buffer because it greatly affects the HT voltage. Purified protein samples had 200 mM NaCl in the buffer, and a Vivaspin 50 kDa cutoff spin concentrator was used to buffer exchange the protein into 5 mM Tris pH 7.5, 1 mM DTT, 5 μM PLP.

CD measurements were conducted at two different concentrations His-tagged AK:DapDc: 0.2 mg/mL and 0.4 mg/mL. These samples of AK:DapDc are not 100% pure. The spectra of the two concentrations show a high degree of correlation and indicate that most of the secondary structures present are β-elements (Figure 5.1). There is a small peak at 220 nm suggesting there are some α-helical elements present also.

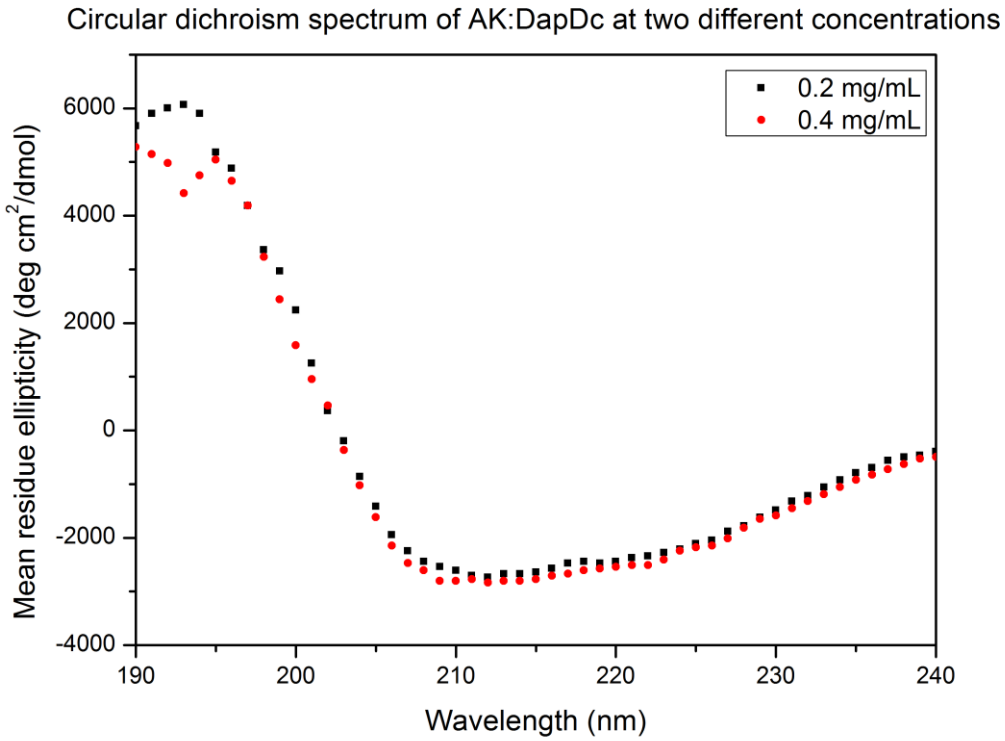


Figure 5.1 - Circular dichroism spectrum of AK:DapDc at 0.2 mg/mL and 0.4 mg/mL.

Online CD analysis tools CDNN, K2D, and ContinLL (via Dichroweb, http://dichroweb.cryst.bbk.ac.uk/) were used to determine the proportions of secondary structures from the CD spectra of two different concentrations of AK:DapDc (Table 5.1)[10-13]. The proportions of secondary structures predicted by PSIPRED and homology modelling (See section 5.3) are different to the observed proportions. This could be due to the protein being partially unfolded, the sample not being very pure, or the prediction being incorrect. The CD profile of an unfolded protein has predominantly β-elements and few α-helices[1]. DSF tests showed that with all samples there was some degree of fluorescence present before the thermal shift assay. It is likely that AK:DapDc is partially denatured and unfolding at 20°C.

Table 5.1 - Proportions of secondary structures present in AK:DapDc.

Secondary elements	PSIPRED prediction	CDNN 0.2 mg/mL	CDNN 0.4 mg/mL	K2D	ContinLL 0.4 mg/mL
α-helix	36.1%	16 %	16 %	9 %	11 %
β-strands	21.5%	50 %	50 %	43 %	57 %
Random coil	42.4%	34 %	34 %	48 %	32 %

5.3 Homology Models

To predict a possible three-dimensional structure for AK:DapDc, homology models were created with SWISS-MODEL (http://swissmodel.expasy.org/)[4, 14, 15]. The AK:DapDc translated amino acid sequence was entered into SWISS-MODEL. The bifunctional AK:DapDc has the highest sequence similarity with two proteins: the DapDc domain has highest similarity with the DapDc from *E. coli* (*Ec*DapDc - 1KNW) while the AK domain has highest similarity with the AK from *Arabidopsis thaliana* (*At*AK – 2CDQ).

5.3.1 DapDc – 1KNW

The C-terminal domain of AK:DapDc (residues 510-867) has the highest similarity with the *E. coli* DapDc (1KNW). 1KNW is 425 residues long and contains 30% α-helix, 22% β-sheet, and 48% random coil. This model template had ~30% identity with AK:DapDc (Figure 5.2).

All DapDc enzymes are TIM-barrel enzymes with an additional C-terminal β-sandwich domain and belong to the fold type III class of PLP-dependent enzymes[16]. This type of PLP-dependent enzyme is catalytically active as a dimer, as both monomers contribute residues to the active site[17]. DapDc is a homodimer and the active sites are formed by the

interface of both subunits[18]. This suggests that AK:DapDc is also an obligate dimer in order to be catalytically active.

DapDc requires pyridoxal 5-phosphate (PLP) as a cofactor to carry out its reaction. SWISS-MODEL modelled PLP into the DapDc homology model, as all of the residues that interact with PLP in the template (1KNW) are present in AK:DapDc (Figure 5.2). PLP is located in the active site of DapDc at the C-terminal face of the TIM barrel region.

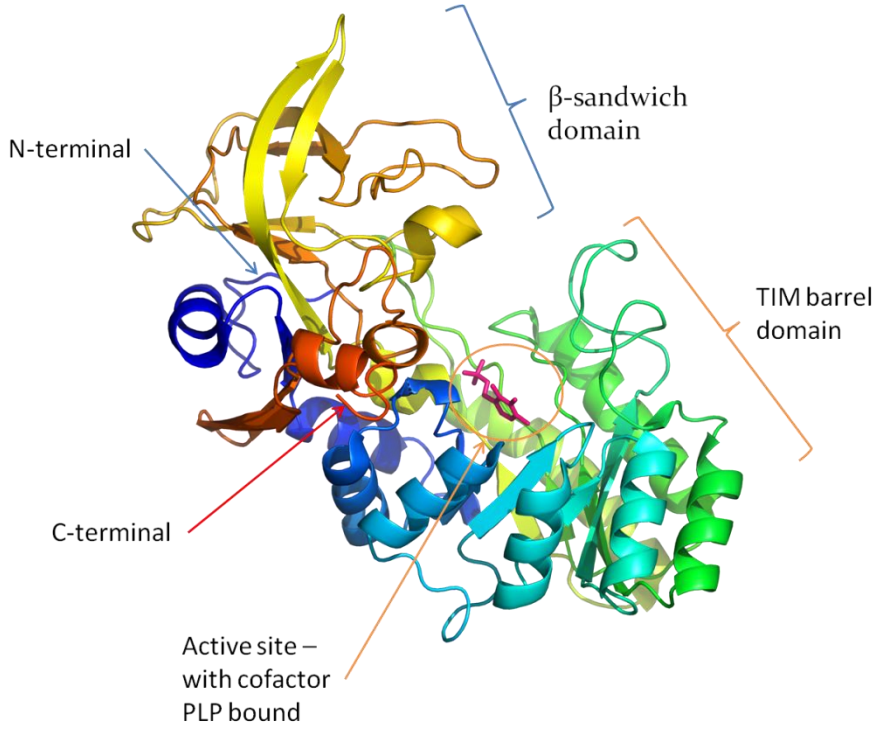


Figure 5.2. Homology model of *Xf*DapDc based on the structure of *Ec*DapDc (PDB ID: 1KNW)

5.3.2 AK – 2CDQ

Homology modelling of the AK domain of the AK:DapDc amino acid sequence (residues 12-480) using SWISS-MODEL predicts the domain to have a structure similar to that of the AK from *Arabidopsis thaliana* (PDB ID: 2CDQ)[19]. The sequence identity was ~30%. This is also very similar to the structure of *E.coli's* AK III, which forms dimers and

tetramers in solution (also ~30% identity). This structure does not have any substrates or cofactors modelled in (Figure 5.3).

AK has two domains: an N-terminal catalytic domain, and a C-terminal regulatory domain (Figure 5.3). The catalytic domain has two lobes: the N-lobe and the C-lobe. The active site of the enzyme lies between the two lobes of the catalytic domain. The regulatory domain also has two parts: two ACT (aspartate kinase:chorismate mutase:TyrA) regulatory domain repeats. ACT regulatory domains have a conserved βαββαβ topology and are often found in pairs, i.e. tandem repeats[20].

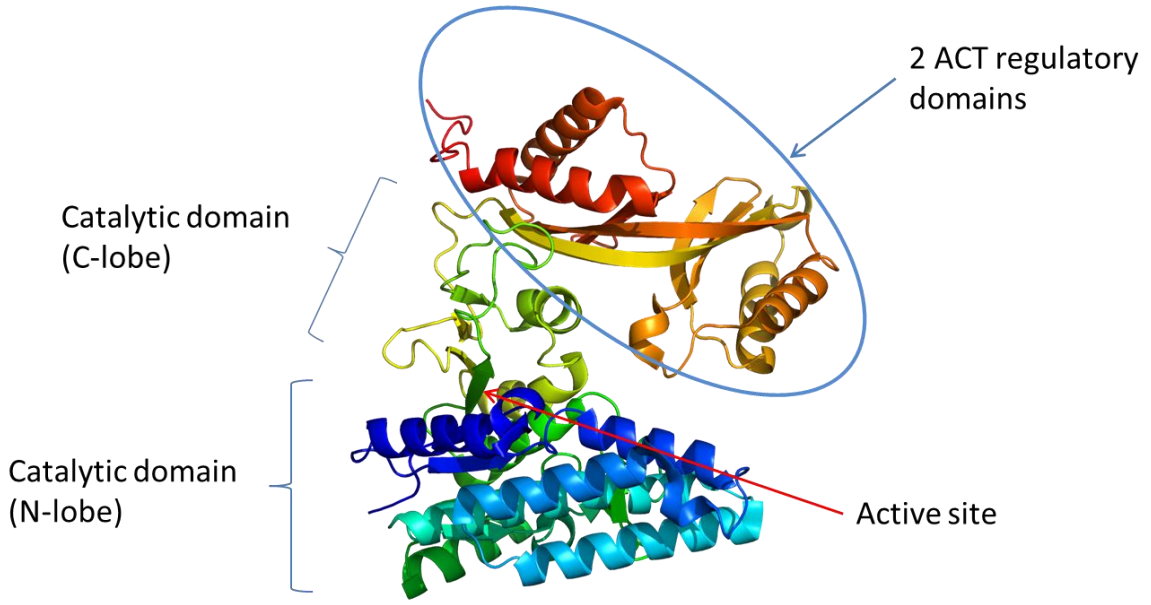


Figure 5.3. Homology model of AK from *Arabidopsis thaliana* (PDB ID: 2CDQ)

AK enzymes can dimerise in two different ways: either the regulatory domains of each monomer associate or the two catalytic domains interact[21]. The *E.coli* AK dimerises by association of the regulatory ACT domains[22]. It also forms a tetramer, by association of the catalytic domains of each dimer[22]. This suggests that the interface region required for tetramerisation in *Ec*AK may be present in AK:DapDc and facilitate dimerisation.

5.4 Analytical Ultracentrifugation (AUC)

The purpose of AUC in this study was to determine the molecular weight and oligomeric state of AK:DapDc in solution. The first sample of AK:DapDc analysed by AUC was His-tagged, but purified by anion exchange chromatography and size exclusion and eluted as one peak (See section 3.4).

Three concentrations of AK:DapDc were used to investigate the concentration dependence of any higher oligomeric states: 0.1 mg/mL, 0.4 mg/mL, 0.8 mg/mL. Wavelength and radial scans determined 290 nm and 5.8-7.3 cm radial length to be good parameters for the experiment. The samples were spun at 40,000 rpm and 250 scans were taken. The Fit data and residuals plots for each concentration are shown in Figure 5.5 and Figure 5.6.

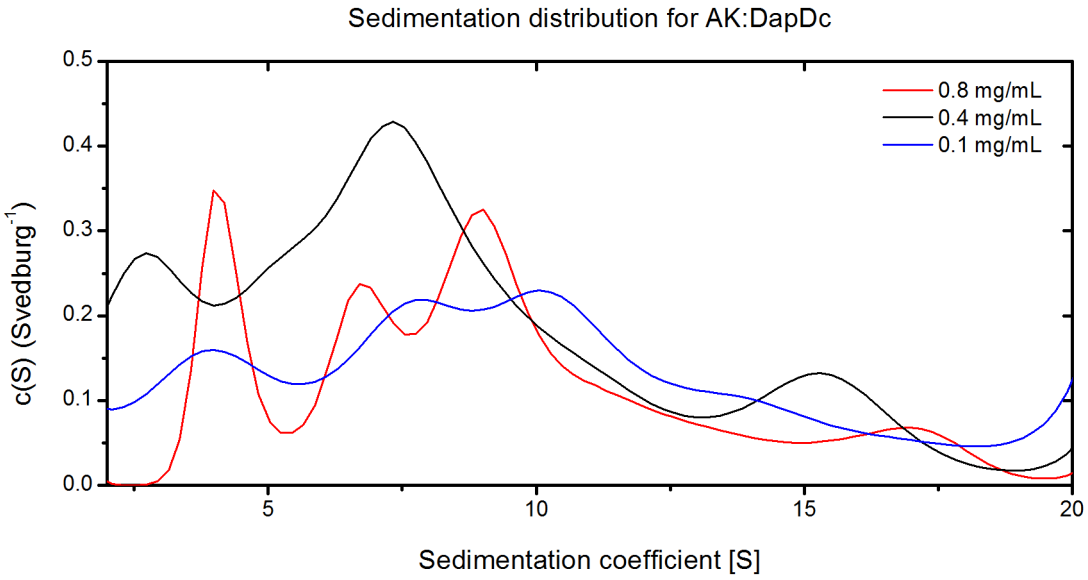
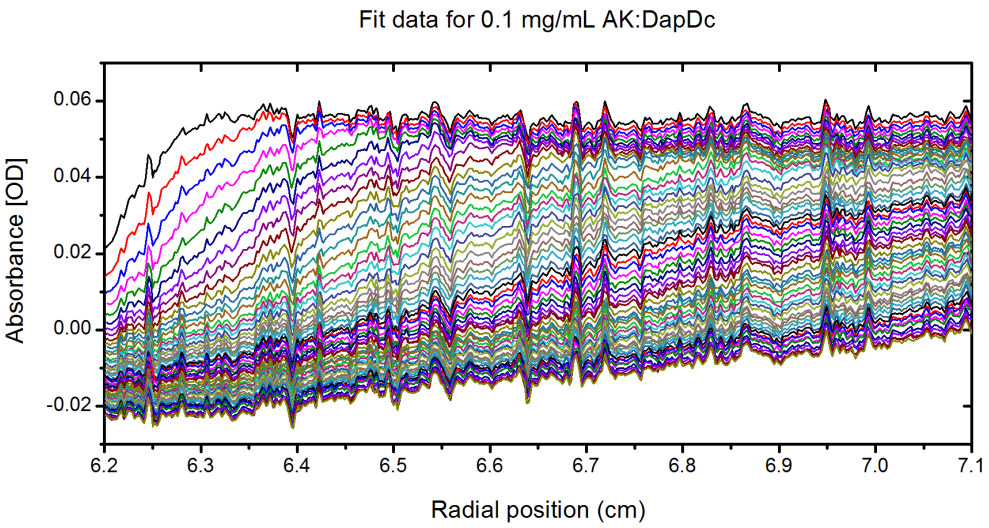


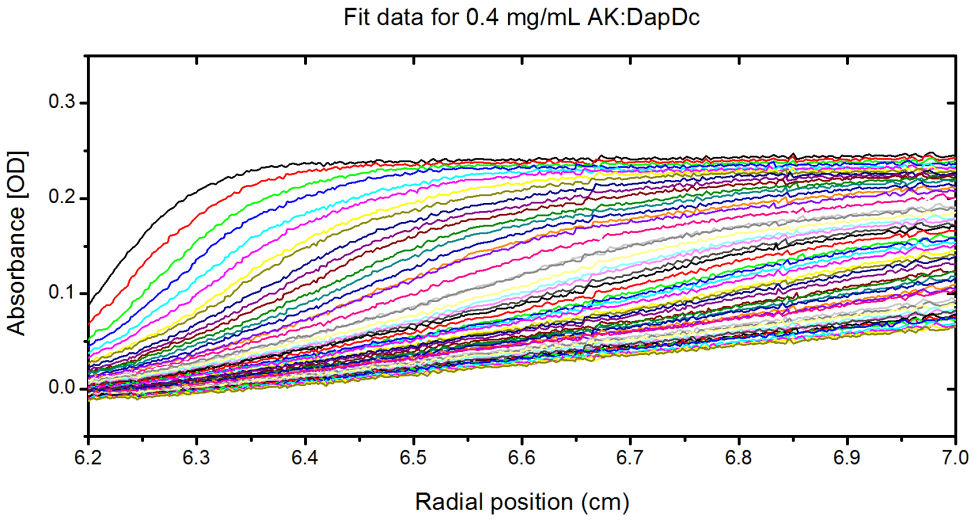
Figure 5.4. Sedimentation distribution of AK:DapDc purified by anion exchange and size exclusion, analysed in sedfit. Normalized to (0 1.0)

The sedimentation coefficient distribution plots for the three concentrations are overlaid in Figure 5.4 to investigate whether there is a concentration-dependent dimerization of AK:DapDc. The sample of the lowest concentration (0.1 mg/mL) has a peak at ~4.2 [S] corresponding to ~88.5 kDa monomeric AK:DapDc and a peak at ~6.8 [S] corresponding to ~180 kDa dimeric AK:DapDc. These proteins are 15.6% and 19.4% of the total protein present, respectively. The sample of 0.4 mg/mL AK:DapDc has a peak at ~9 [S] corresponding to ~164 kDa dimeric AK:DapDc. This protein makes up 56.9% of the total protein present. The sample with the highest concentration (0.8 mg/mL) has a peak at ~6.8 [S] corresponding to ~96.8 kDa monomeric AK:DapDc, and a peak at ~10 [S] corresponding to ~179 kDa dimeric AK:DapDc. These two peaks are 13.7% and 37.5% of the total protein. Comparing the lowest and highest concentration data, there is a much higher proportion of dimer in the 0.8 mg/mL sample and a more equal proportion of monomer and dimer at the lowest concentration. This suggests that the higher the concentration of AK:DapDc, the greater the proportion of dimer present. However, even at low concentrations there is still a detectable amount of dimer present.

A.



B.



C.

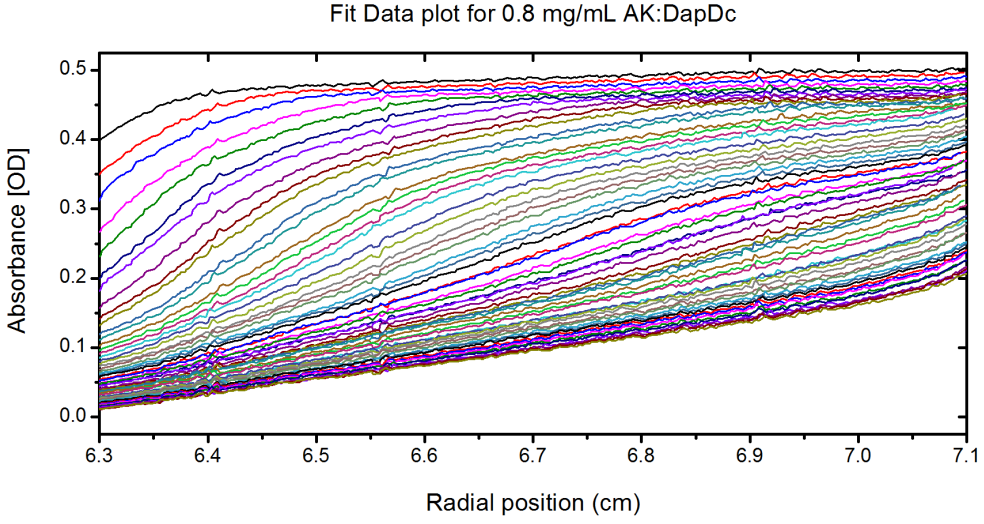
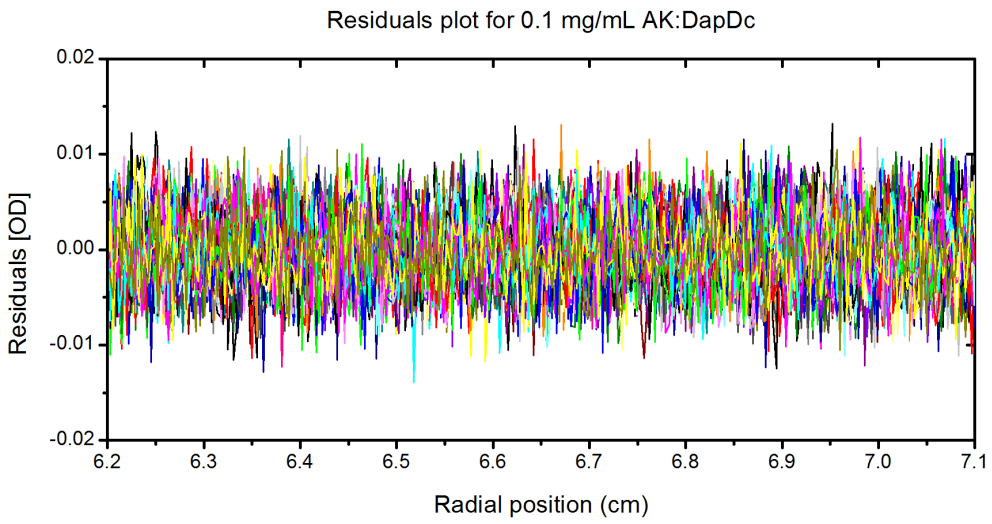
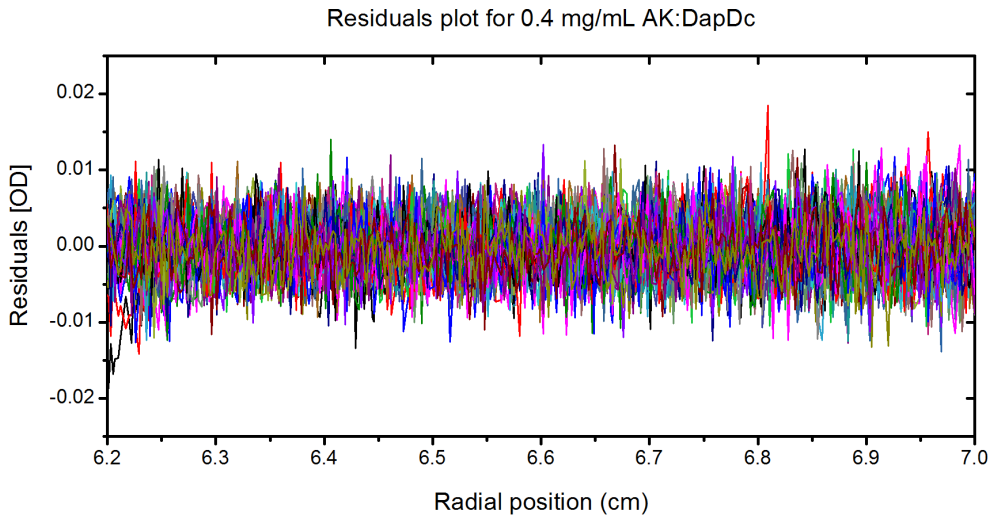


Figure 5.5. AUC fit data for AK:DapDc sample purified by anion exchange and size exclusion, analysed in sedfit.

A.



B.



C.

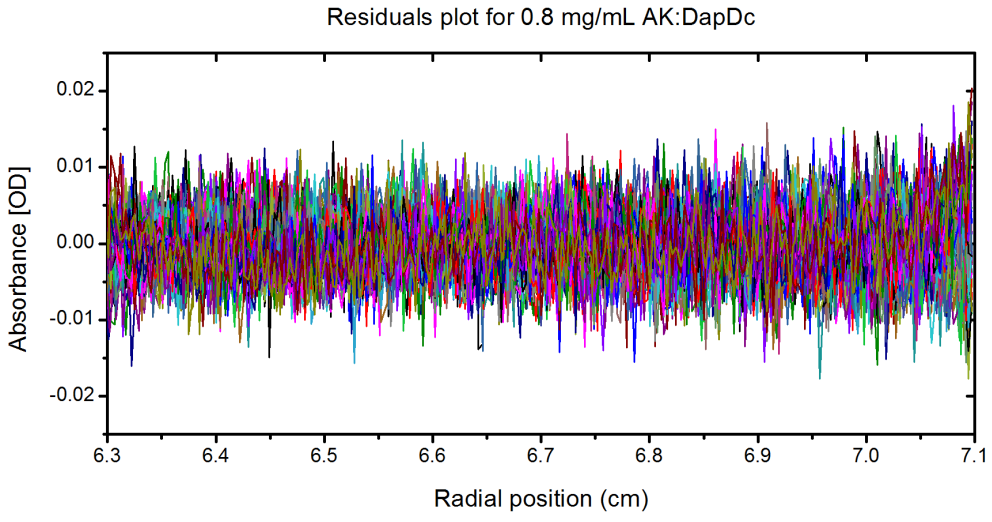


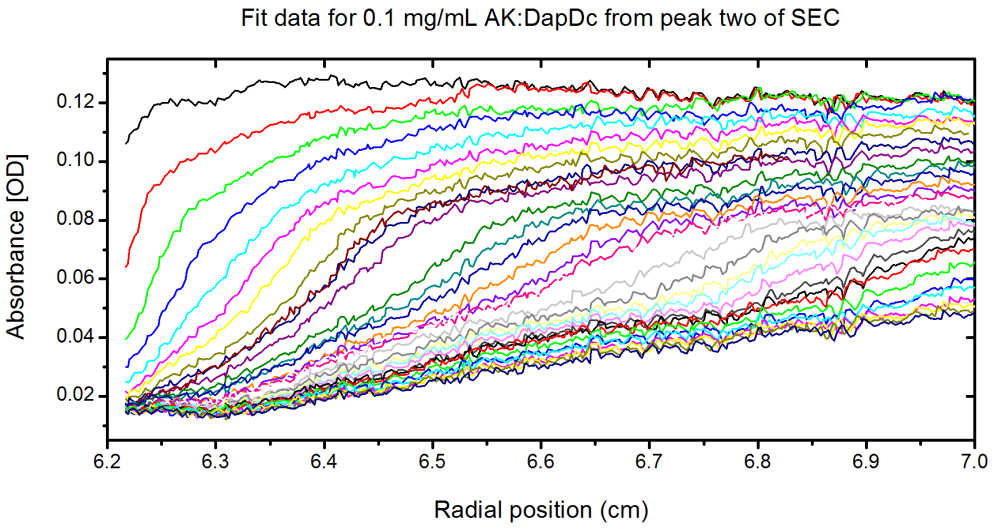
Figure 5.6. AUC residuals data for AK:DapDc purified by anion exchange and size exclusion, analysed in sedfit.

The second set of AK:DapDc samples analysed by AUC was His-tagged and purified by His-Trap chromatography and size exclusion chromatography (SEC). AK:DapDc eluted from SEC as three peaks. The first peak eluted within the void volume of the column and was determined by DSF to be aggregated. Peaks two and three had a small amount of aggregation, but were folded and DSF gave a melting profile for both peaks. Three concentrations of peak two were analysed by AUC: 0.1 mg/mL, 0.4 mg/mL, and 0.8 mg/mL. Two concentrations of peak three were analysed: 0.1 mg/mL and 0.8 mg/mL.

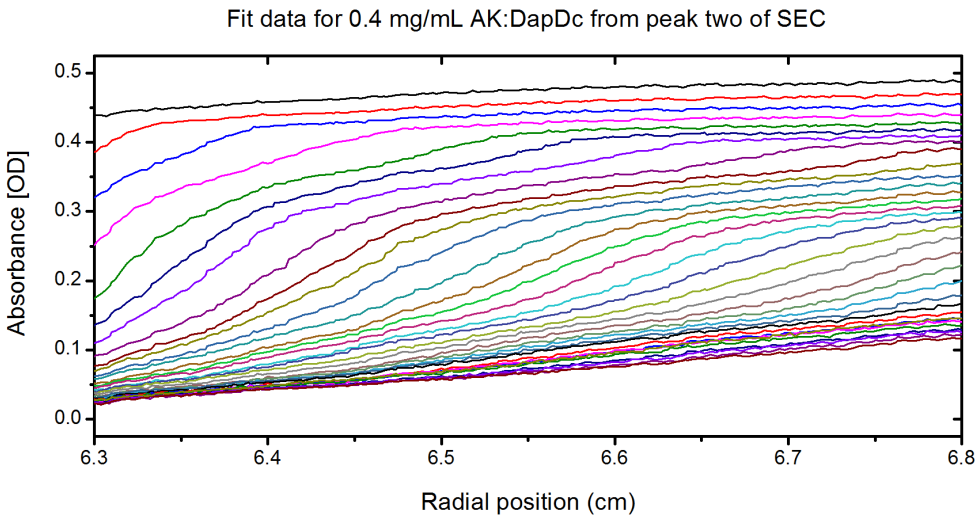
Peak two from SEC was analysed at 288 nm and at 40,000 rpm. 150 scans were taken. Fit data and residuals plots for each concentration are shown in Figure 5.7 and Figure 5.8.

The sample with the lowest concentration (0.1 mg/mL) resulted in a sedimentation distribution with a peak at ~7.4 [S] corresponding to ~97.8 kDa monomeric AK:DapDc (Figure 5.9). This protein constitutes 24.4% of the total protein in the sample. The sample of 0.4 mg/mL AK:DapDc had a peak at ~8.99 [S] potentially corresponding to ~160 kDa dimeric AK:DapDc. This protein is 17.5% of the protein in the sample. The most concentrated AK:DapDc sample (0.8 mg/mL) had a distribution profile with a peak at ~6.5 [S] which corresponds to ~99.4 kDa monomeric AK:DapDc and a peak at ~10.3 [S] which indicates the presence of ~198 kDa dimeric AK:DapDc. These proteins constitute 17.2% and 11.8% of the total protein sample, respectively. There are predominantly proteins corresponding to the predicted size of an AK:DapDc dimer (~195 kDa). This suggests that this peak largely contains dimeric AK:DapDc.

A.



B.



C.

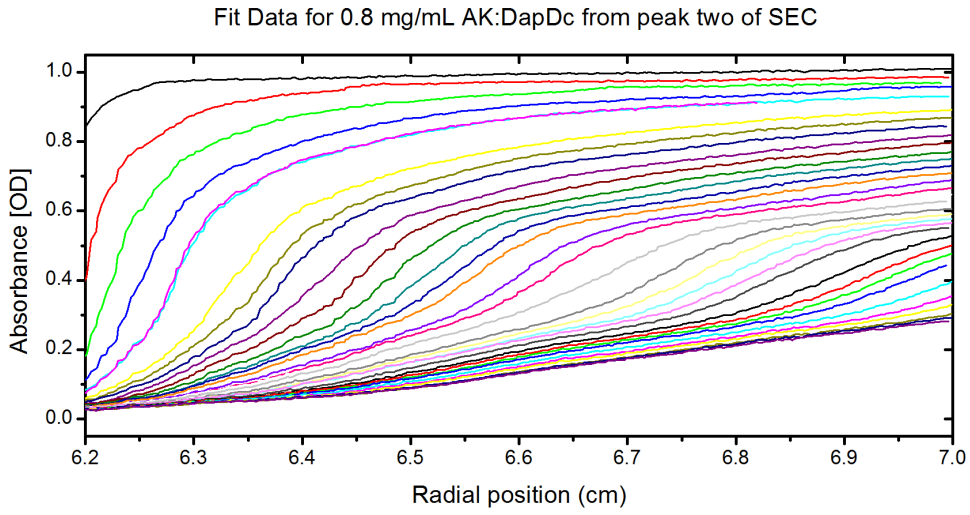
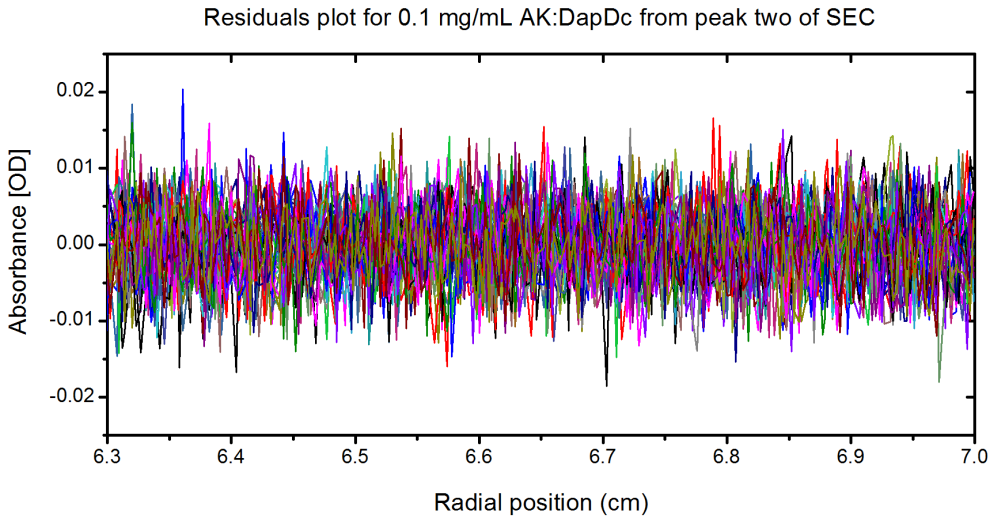
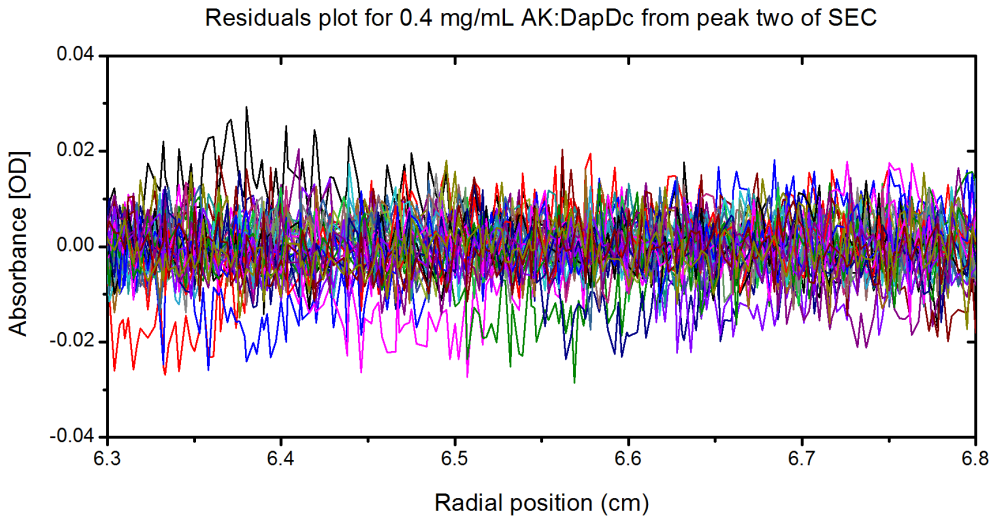


Figure 5.7. AUC fit data plots for peak two of AK:DapDc purified by His-Trap and size exclusion, analysed in sedfit.

A.



B.



C.

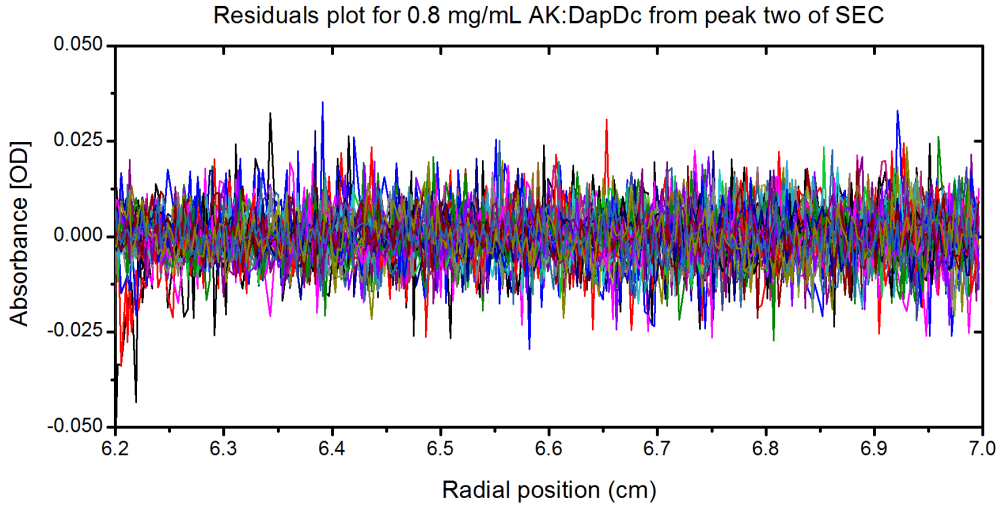


Figure 5.8. AUC residuals plots for peak two of AK:DapDc purified by His-Trap and size exclusion, analysed by sedfit.

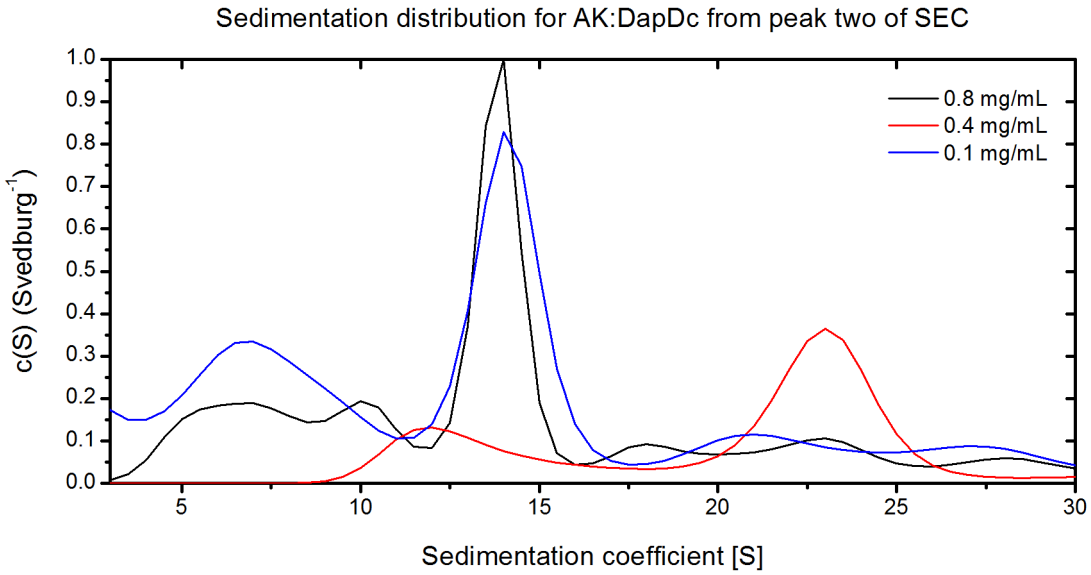
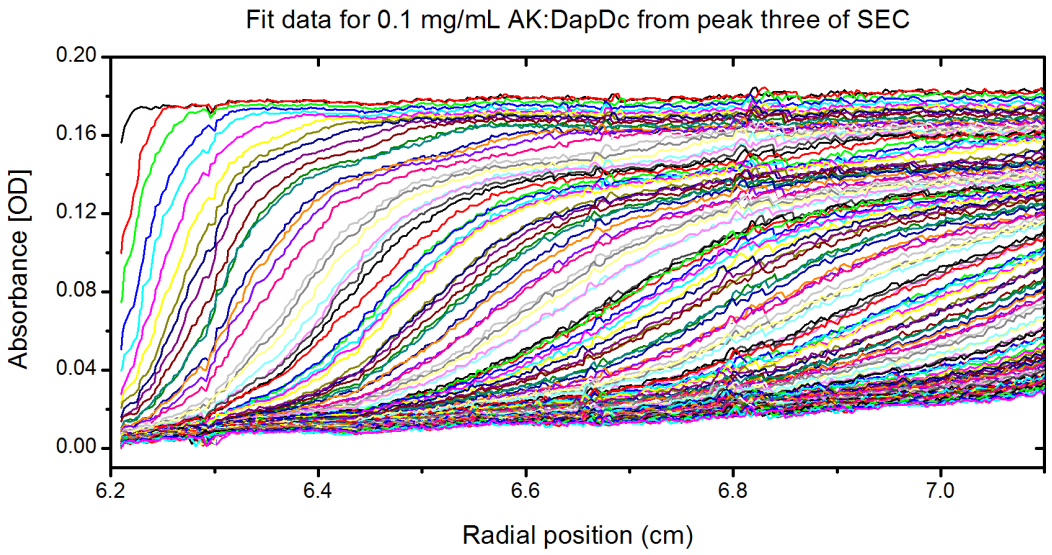


Figure 5.9. Sedimentation distribution for AK:DapDc from peak two of size exclusion, analysed in sedfit. Normalized to (0-1.0)

Peak three from AUC was analysed at 285 nm, 45,000 rpm and at two concentrations (0.1 mg/mL and 0.8 mg/mL). The fit data and residuals plots for each concentration are shown in Figure 5.10 and Figure 5.11. This peak from SEC is hypothesised to be monomeric AK:DapDc. The lower concentration (0.1 mg/mL) produced a sedimentation distribution with a peak at ~6.2 [S] corresponding ~102 kDa monomeric AK:DapDc (Figure 5.12). This peak constitutes 74.2% of the protein in the sample. There was no peak corresponding to dimeric AK:DapDc (~195 kDa). The sedimentation distribution for the higher concentration of AK:DapDc (0.8 mg/mL) has a peak at ~6.2 [S] corresponding to monomeric AK:DapDc. This protein is 68.5% of the protein in the sample. This concentration shows the presence of larger proteins more than the 0.1 mg/mL sample, but there is no peak corresponding to dimeric AK:DapDc. This suggests that the size exclusion peak three is indeed monomeric AK:DapDc, and that this state is not concentration dependent in the range measured.

A.



B.

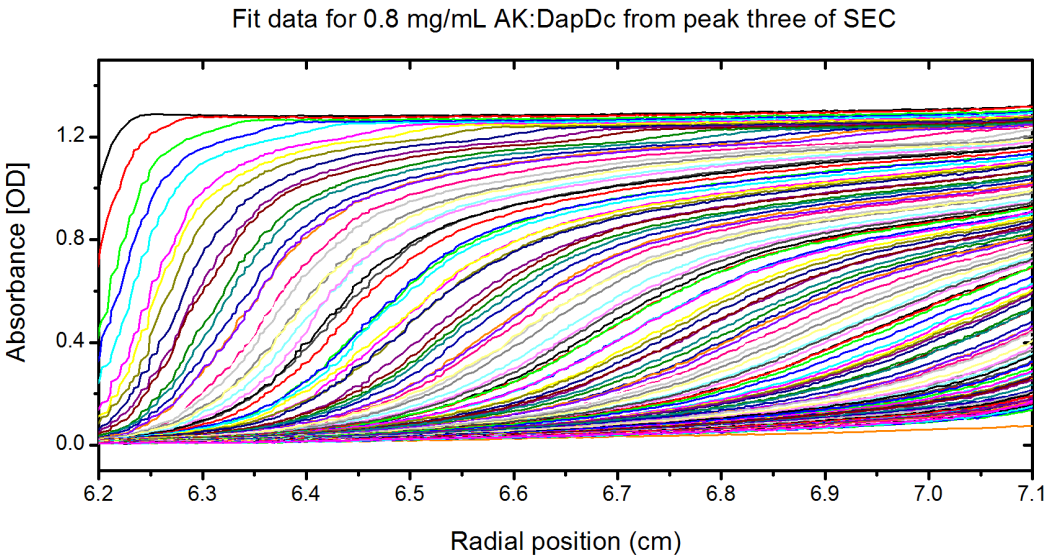
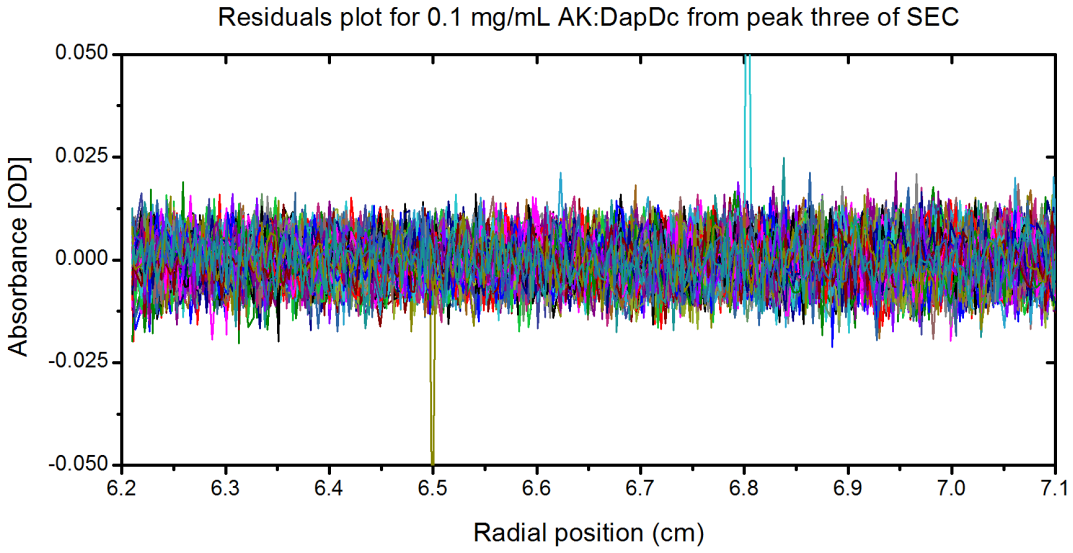


Figure 5.10. AUC fit data for peak three of AK:DapDc purified by His-Trap and size exclusion, and analysed by sedfit.

A.



B.

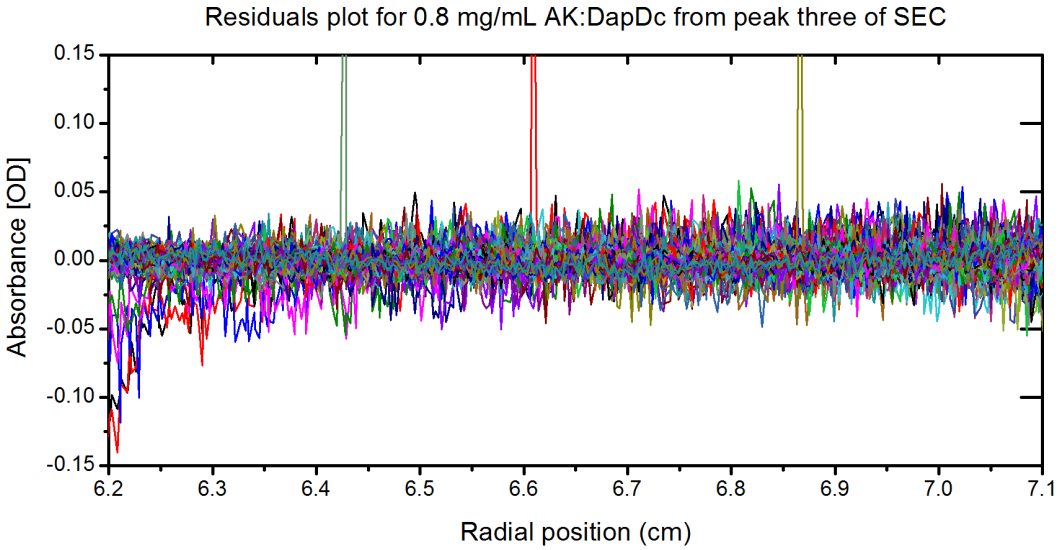


Figure 5.11. AUC residuals plots for peak three of AK:DapDc purified by His-Trap and size exclusion, analysed by sedfit.

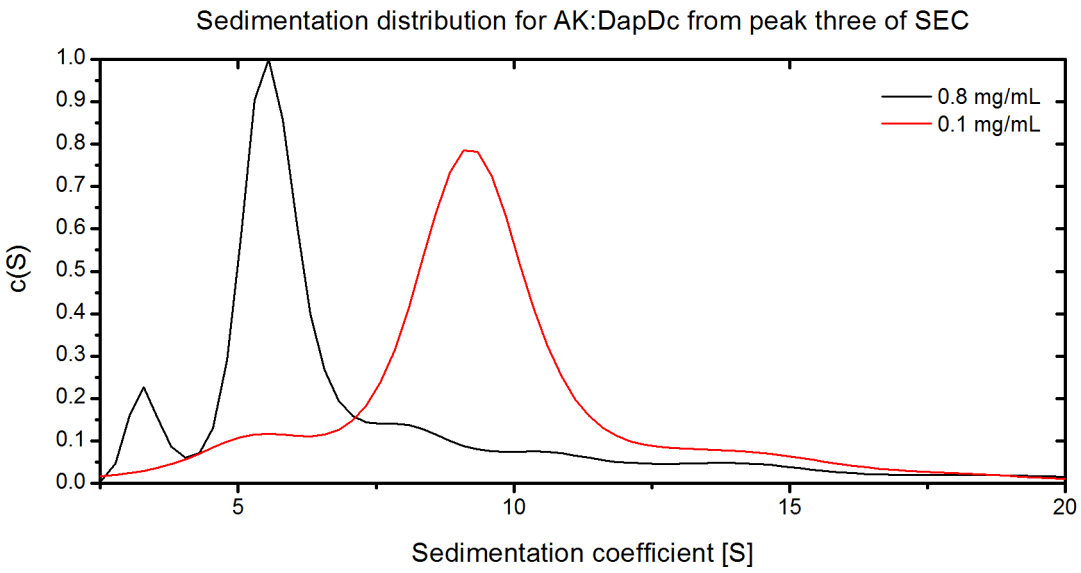


Figure 5.12. Sedimentation distribution for AK:DapDc from peak three of size exclusion, analysed by sedfit. Normalized to (0-1.0)

5.5 Crystal Screens

Four crystal screens were trialled at two temperatures (5°C and 20°C). The protein used in this trial was concentrated using a Vivaspin spin concentrator (Sartorius). The four crystals screens are: Clear Strategy Screen™ I HT-96, JCSG-*plus*™ HT-96, Morpheus™, and PACT *premier*™ HT-96 (moleculardimensions.com).

Clear Strategy Screen™ I HT-96 screens from pH 5.5 – 8.5, combining different PEGs (precipitants) and salts, to give an easily optimised, high-throughput screen. JCSG-*plus*™ HT-96 is a sparse-matrix screen based on previously successful conditions. It screens from pH 4.0-10.0, using neutralized organic acids and polyalcohol conditions. Morpheus™ covers a range of pH (6.5-8.5), PEGs, and salts, but also includes low molecular weight ligands found to occur in many crystal structures. PACT *premier*™ HT-96 is a pH (4.0-9.0), anion, cation, trial, which tests each of these factors against the precipitant PEG 3350.

No crystals were found in any of the trials. Each screen trialled at both 5°C and 20°C to test whether the protein was more stable at a low temperature and if this would slow down the process of concentration to give crystals rather than precipitate. In both temperatures, only precipitate or amorphous structures were found (Figure 5.13). DSF showed that AK:DapDc was denatured at pH 4-6, and was folded at pH 6 but not particularly stable. Buffers that caused precipitation and denaturation during buffer exchange and DSF (e.g. HEPES), resulted in precipitation instead of crystallisation. Magnesium chloride does not stabilise AK:DapDc, thus crystallisation conditions containing magnesium chloride also resulted in precipitate.

A. B.

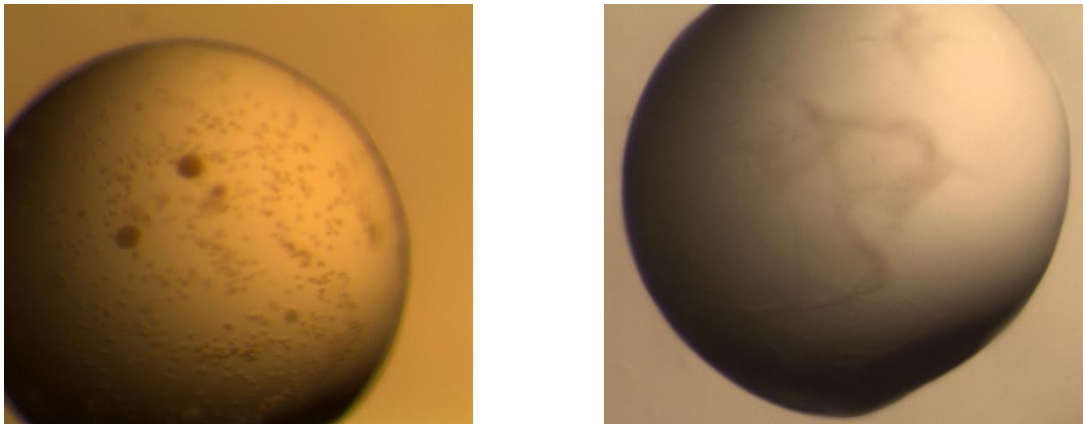


Figure 5.13. Examples of precipitate (A.) and amorphous structures (B.) found in all crystal screens.

Morpheus™ uses amino acids such as racemic lysine to screen ligands that are likely to co-crystallize with a protein. Lysine is the predicted inhibitor for both domains of AK:DapDc, and so binding of lysine to either or both domains may stabilise AK:DapDc and aid in crystallisation. It is unclear whether it will aid in crystallisation as the combination of conditions resulted in precipitation, despite the presence of lysine.

Ammonium sulphate precipitates AK:DapDc (See Section 3.4.1), thus the conditions in JCSG-plus HT-96, Morpheus™, and PACT premier HT-96 containing above 0.1M ammonium sulphate have resulted in precipitation of AK:DapDc.

5.6 Summary of Results

CD analysis of AK:DapDc at two concentrations estimated significantly different proportions of 2° structures present compared to computational prediction and homology modelling. The proportion of α-helical elements present is much lower, indicating that AK:DapDc is unstable and unfolding at 20°C, or the predicted structure of AK:DapDc is very different from the actual structure in solution. It is more likely that AK:DapDc is unfolded and partially denatured, as it is an unstable protein (See Chapter three).

Homology modelling created two models for AK:DapDc from the monofunctional forms of aspartate kinase and diaminopimelate decarboxylase from Arabidopsis thaliana and Escherichia coli, respectively. These two template enzymes are well documented and so information about the domain structure and oligomeric state of AK:DapDc can be inferred. This model will need to be confirmed by x-ray crystallography. There is a linker connecting the two functions of AK:DapDc which has not been modelled. This linker is approximately 30 residues long and its flexibility may be affecting the 3° and 4° structure of AK:DapDc. This could explain the difference between the CD results and the predicted secondary structure.

The 3° and 4° structure of purified AK:DapDc was investigated using AUC. AK:DapDc samples purified using anion exchange and SEC contain both monomers and dimers of AK:DapDc, with the higher concentration containing a much greater proportion of dimer. This may be due to the prediction that AK:DapDc is an obligate dimer for DapDc to be

active. AK:DapDc purified by IMAC and SEC eluted from SEC in three peaks. Peaks two and three were hypothesised to contain AK:DapDc dimer and monomer, respectively. AUC results suggest that this may be the case.

References

1. Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure, *Nature Protocols 1*, 2876-2890.
2. Buchan, D. W. A., Minneci, F., Nugent, T. C. O., Bryson, K., and Jones, D. T. (2013) Scalable web services for the PSIPRED Protein Analysis Workbench, *Nucleic Acids Research 41*, W349-W357.
3. Bordoli, L., and Schwede, T. (2012) Automated Protein Structure Modeling with SWISS-MODEL Workspace and the Protein Model Portal, In *Homology Modeling* (Orry, A. J. W., and Abagyan, R., Eds.), pp 107-136, Humana Press.
4. Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Cassarino, T. G., Bertoni, M., Bordoli, L., and Schwede, T. (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, *Nucleic Acids Research 42*, W252-W258.
5. Ghirlando, R. (2011) The analysis of macromolecular interactions by sedimentation equilibrium, *Methods 54*, 145-156.
6. Cole, J. L., Lary Jw Fau - P Moody, T., P Moody T Fau - Laue, T. M., and Laue, T. M. Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium.
7. Lebowitz, J., Lewis, M. S., and Schuck, P. (2002) Modern analytical ultracentrifugation in protein science: A tutorial review, *Protein Science 11*, 2067-2079.
8. Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, *Nature Protocols 1*, 2527-2535.
9. Greenfield, N. J. (2006) Analysis of the kinetics of folding of proteins and peptides using circular dichroism, *Nature Protocols 1*, 2891-2899.
10. Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) EVALUATION OF SECONDARY STRUCTURE OF PROTEINS FROM UV CIRCULAR-DICHROISM SPECTRA USING AN UNSUPERVISED LEARNING NEURAL-NETWORK, *Protein Engineering 6*, 383-390.
11. Bohm, G., Muhr, R., and Jaenicke, R. (1992) QUANTITATIVE-ANALYSIS OF PROTEIN FAR UV CIRCULAR-DICHROISM SPECTRA BY NEURAL NETWORKS, *Protein Engineering 5*, 191-195.
12. Provencher, S. W., and Glockner, J. (1981) ESTIMATION OF GLOBULAR PROTEIN SECONDARY STRUCTURE FROM CIRCULAR-DICHROISM, *Biochemistry 20*, 33-37.
13. Lobley, A., Whitmore, L., and Wallace, B. A. (2002) DICHROWEB: an interactive website for the analysis of protein secondary structure from circular dichroism spectra, *Bioinformatics 18*, 211-212.
14. Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, *Bioinformatics 22*, 195-201.
15. Kiefer, F., Arnold, K., Künzli, M., Bordoli, L., and Schwede, T. (2009) The SWISS-MODEL Repository and associated resources, *Nucleic Acids Research 37*, D387-D392.

16. Ray, S. S., Bonanno, J. B., Rajashankar, K. R., Pinho, M. G., He, G. S., De Lencastre, H., Tomasz, A., and Burley, S. K. (2002) Cocrystal structures of diaminopimelate decarboxylase: Mechanism, evolution, and inhibition of an antibiotic resistance accessory factor, *Structure 10*, 1499-1508.
17. Eliot, A. C., and Kirsch, J. F. (2004) Pyridoxal phosphate enzymes: Mechanistic, structural, and evolutionary considerations, *Annual Review of Biochemistry 73*, 383-415.
18. Kefala, G., Perry, L. J., and Weiss, M. S. (2005) Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of LysA (Rv1293) from Mycobacterium tuberculosis, *Acta Crystallographica Section F-Structural Biology and Crystallization Communications 61*, 782-784.
19. Mas-Droux, C., Curien, G., Robert-Genthon, M., Laurencin, M., Ferrer, J.-L., and Dumas, R. (2006) A Novel Organization of ACT Domains in Allosteric Enzymes Revealed by the Crystal Structure of Arabidopsis Aspartate Kinase, *The Plant Cell Online 18*, 1681-1692.
20. Chipman, D. M., and Shaanan, B. (2001) The ACT domain family, *Current Opinion in Structural Biology 11*, 694-700.
21. Robin, A. Y., Cobessi, D., Curien, G., Robert-Genthon, M., Ferrer, J. L., and Dumas, R. (2010) A New Mode of Dimerization of Allosteric Enzymes with ACT Domains Revealed by the Crystal Structure of the Aspartate Kinase from Cyanobacteria, *Journal of Molecular Biology 399*, 283-293.
22. Kotaka, M., Ren, J., Lockyer, M., Hawkins, A. R., and Stammers, D. K. (2006) Structures of R- and T-state Escherichia coli aspartokinase III - Mechanisms of the allosteric transition and inhibition by lysine, *Journal of Biological Chemistry 281*, 31544-31552.

6. Discussion

6.1 Introduction

Xylella fastidiosa is greatly researched worldwide due to its significant impact on many economically important crops. *X. fastidiosa* has not yet been detected in New Zealand, but overseas evidence has shown that it affects commercial crops known to be important to the New Zealand economy[1, 2]. The New Zealand wine industry is worth $1.5 billion dollars annually, and exports $1.2 billion dollars of wine annually[3]. The ornamental plant and floriculture industry is worth approximately $70 million annually[4]. Biosecurity New Zealand considers *X. fastidiosa* to be enough of a risk that it has developed an assay for the detection of *X. fastidiosa* [5].

Research into its biology and mechanism of pathogenesis are essential for combating *X. fastidiosa* infection, more so because the bacteria can remain dormant in grapes throughout much of the year and can spread to over 100 species of plants where it can also remain dormant and asymptomatic[6]. *X. fastidiosa*'s metabolism and biosynthetic pathways are equally as important to study as many of these are essential pathways.

Aspartate kinase:Diaminopimelate decarboxylase (AK:DapDc) is the focus of this study because it is part of an essential pathway for *X. fastidiosa*, and is an intriguing and extremely rare bifunctional enzyme. Due to its uniqueness, this enzyme potentially offers an antibiotic target to control *X. fastidiosa* infections[7].

6.2 Expression of AK:DapDc

Over-expression of Aspartate kinase:Diaminopimelate decarboxylase was achieved in this study, after many expression trials. Although over-expression of AK:DapDc can be achieved in *Escherichia coli* BL21 cells, upon cell lysis and purification most of the AK:DapDc protein is insoluble and aggregated. Buffer screens and trials have improved AK:DapDc solubility and stability enough to perform a number of biophysical techniques on the protein. Other techniques that would further characterise the protein and support the data and hypotheses in this study require more pure, stable purified protein. If the stability issues arise from mis-folding, which facilitates aggregation, then the most effective solution may be to modify the expression system and conditions.

E. coli has many advantages as an expression system, but not every gene can be expressed well in this organism [8]. A significant fraction of some recombinant proteins fails to fold correctly and aggregates into inclusion bodies [9]. This can be due to properties of the gene sequence such as differences in codon usage between the recombinant gene and *E. coli* cells[10], the protein's ability to fold correctly, the stress of over-expression on the cells, or suboptimal intracellular conditions (e.g. pH or redox conditions) [11]. Codon usage was investigated using Graphical Codon Usage Analyser (www.gcua.de) which compares codon usage in a query gene sequence to the codon usage of another species, e.g. *E. coli*. There were some differences in codon usage, but the gene overexpressed well in *E. coli* so optimisation was not carried out.

This study has trialled alternative expressions systems and conditions such as expression in *E. coli* BL21 DE3 Rosetta cells which contain rare tRNA genes, expression with a GST-tag to aid solubility and folding, media supplementation with 1% glucose, short expression time with high level IPTG induction, slow expression at 4°C with low IPTG induction.

AK:DapDc expression was trialled in *E. coli* Rosetta cells, but this didn't appear to significantly improve the solubility of the recombinant protein. Codon usage in the AK:DapDc gene (XFLM_07435) was compared to that used by *E. coli*, and there were a few noticeable differences. The effect of expression with extra tRNA genes may not be enough to help correct the mis-folding of AK:DapDc.

Cloning the AK:DapDc gene (XFLM_07435) into a vector (pET30-GST) so that there is an N-terminal GST tag has been attempted in this study. The GST tag is both a solubility and affinity tag, and aids in the stability of proteins it is fused to. Despite sequencing confirming that the gene was inserted in the vector, a GST-AK:DapDc fusion protein did not express.

The addition of 1% glucose to pre-cultures and large cultures was also trialled in this study. The addition of glucose prevents leaky expression[12] and is also reported to increase the solubility of recombinant proteins. The glucose did not significantly improve the solubility of AK:DapDc. Likewise, varying the length of time given for expression and the concentration of IPTG did not have a noticeable effect on solubility.

Other promising methods that have not been trialled include auto-inducing medium, co-expression with chaperone proteins that aid in proper folding and separate expression of the AK:DapDc domains.

Auto-inducing medium is one method used to improve the solubility of recombinant proteins[13]. The medium contains lactose, which *E. coli* breaks down into galactose and glucose. The glucose is used as an energy source and also inhibits induction. The galactose induces expression of recombinant genes inserted into vectors which contain a T7 lac promoter. The slow break down of lactose restricts the amount of galactose that is available

to induce expression and thus expression occurs slowly. Slower expression can aid in better folding of some recombinant proteins. Media additives have improved the solubility of small proteins (25-45 kDa).

While small proteins are often helped by the conditions already mentioned large recombinant proteins may require more assistance. Chaperone co-expression or expression as two separate domains may have the greatest effect on AK:DapDc solubility.

E. coli has endogenous chaperones which assist in folding at least 300 *E. coli* proteins [14]. It appears that during over-expression of recombinant protein the cells' chaperones become overwhelmed and so much of the protein becomes mis-folded and aggregates into inclusion bodies [11].

Alternatively, if the overall structure is too unstable, the two halves of AK:DapDc could be expressed separately. Secondary structure prediction can be used to predict the boundaries of secondary structures, and create two conctructs without affecting local secondary structures. A protocol which investigates different combinations of N- and C-terminal positions could be used to identify what length of tail gives the protein the most stability [15]. Separate expression would also simplify enzyme kinetics. When investigating inhibition, expressing the protein as separate domains allows you to tell which domain the inhibitor is binding to, to inhibit activity. It can be established whether or not both domains are inhibited by lysine, and also threonine and methionine. Conversely, it would be intriguing to determine whether binding of an allosteric inhibitor to one domain affects the activity of the other domain.

6.3 Purification of AK:DapDc

Many purification trials were also carried out to try to troubleshoot and optimise the process. AK:DapDc purification was first trialled using Co^{2+} HisTrap chromatography, but the protein would not elute from the column. This problem lead to the hypothesis that AK:DapDc may have been interacting with the Co^{2+} in HisPur column[16]. Aspartate kinase has been shown to interact with cobalt, which becomes oxidised to inert Co^{3+}. This prevents the enzyme from being catalytically active. The reducing agent dithiothreitol (DTT) was recommended for the removal of Co from the metal binding site of aspartate kinase[16]. Adding DTT to the purification buffers allowed the AK:DapDc protein to elute from the Co^{2+} HisTrap column. This was the preferred method of purification in this study as it was much more selective than anion exchange.

The binding of Co^{2+} to the metal binding site of aspartate kinase could be tested by isothermal titration calorimetry (ITC) which measures the heat of interactions of molecules and proteins in solution. This property of cobalt could also be investigated via a SYPRO Orange binding assay which allows detection of nucleotide and divalent cation binding to kinases. The effect of adding Co^{2+} to protein samples could also be analysed by differential scanning fluorimetry (DSF) which would indicate an increase or decrease in AK:DapDc stability with Co^{2+} binding or analytical ultracentrifugation (AUC) which could show a change in overall shape and oligomeric state with Co^{2+} binding.

The benefit of including a reducing agent such as DTT in the purification buffers is that the formation of soluble aggregates may be prevented. There may be unforeseen negative effects of adding a reducing agent. Creating a reducing environment may affect dimerisation and activity.

A major issue in this study was that solubility is not the best marker for quality in purified protein. Mono-dispersity is a more appropriate measure of quality. Soluble aggregates of folded and mis-folded protein have been observed[9], and these may cause further aggregation and instability of properly folded AK:DapDc. They also contaminate purified samples of protein, and may prevent the use of certain techniques on the sample. The use of an aggregation index/rate as a measure of quality would be more effective than merely measuring concentration of soluble protein present. DTT and NaCl were added to the purification buffers in this study, and this was sufficient to prevent further aggregation and allow characterisation techniques to be carried out.

The DSF buffer screen protocol used in this study was very helpful for purification optimisation[17]. The buffer screen was able to identify which buffers, pH range and salt concentration gave the most stable protein (See Section **3.4.1.1**). However, this is effectively a buffer exchange on already purified protein. Whether the buffers that result in the highest melting temperatures are able to solubilise AK:DapDc during cell resuspension and lysis is unsure. Cation exchange chromatography was trialled in this study using two buffers identified by DSF as stabilising AK:DapDc, but was unsuccessful (See Section **3.4.1.1**). This may have been due to the buffers' inability to solubilise AK:DapDc during resuspension and cell lysis. DSF is also useful for detecting aggregation in protein samples, without running a thermal assay. Taking a photo of fluorescence present with the iQ5 RT-PCR machine was used to confirm that SEC peak one consisted of aggregated AK:DapDc (Section **3.4.3**). This showed that peaks two and three, hypothesised to be dimeric AK:DapDc and monomeric AK:DapDc respectively, had a small degree of aggregation/fluorescence.

6.4 Kinetic assay

The parameters calculated in this study for *Sc*SDH are slightly different from those reported in previous studies. In this study SDH was resuspended in 200 mM Tris buffer pH 8.0 and the assay was carried out at 37°C, while Xu *et al*. (2006) used 100 mM HEPES pH 7.0 and the assay was carried out at 25°C[18].

The saccharopine dehydrogenase assay is an effective assay for detecting and measuring the activity of enzymes such as diaminopimelate decarboxylase. It requires only one coupling enzyme and so only one set of kinetic parameters need to be calculated to set up the assay. This is an ideal assay to test for activity in AK:DapDc. However, it cannot be used to test for inhibition of AK:DapDc as the predicted inhibitor, L-lysine[19, 20], is a substrate for *Sc*SDH. In the assay, *Sc*SDH is present in excess, thus any L-lysine present/added is rapidly converted by *Sc*SDH to saccharopine. Thus the activity observed will be *Sc*SDH activity, and not AK:DapDc activity. *Sc*SDH is reported to have absolute specificity for lysine[21] so an L-lysine analogue inhibitor that *Sc*SDH can't convert to saccharopine could be used to investigate inhibition of AK:DapDc. A coupled assay in which the coupling enzyme(s) does not use the inhibitor as a substrate could also be used to test for inhibition, e.g. the PEP carboxylase/malate dehydrogenase coupled assay[22].

6.5 Biophysical characterisation

Circular dichroism (CD), homology modelling and AUC were used to biophysically characterise AK:DapDc. DSF analysis of AK:DapDc also added to the characterisation of the protein. These techniques were used to elucidate information about the secondary, tertiary and quaternary structure of AK:DapDc. CD suggests that AK:DapDc contains both α-helices and β-strands, but the proportions are significantly different from the predicted

secondary structure proportions. AUC and homology modelling indicate that AK:DapDc exists as both monomer and dimer, but may be functional as a dimer. Homology modelling provides clues as to how AK:DapDc might form dimers.

CD gave an average secondary structure profile of AK:DapDc of 13% α-helices, 50% β-strands, and 37% random coil using the programs CDNN[23], K2D[24], and ContinLL[25]. This is significantly different from the proportions predicted computationally: 36% α-helices, 21.5% β-strands, and 42.5% random coil. This could be due to contamination by other proteins present or unfolding of AK:DapDc at 20°C. The purification of AK:DapDc resulted in a protein sample with some small contaminant proteins, which may be altering the secondary structure profile obtained. The CD profile may also be affected by the unfolding of some of the AK:DapDc present. When proteins unfold their secondary structure profile can change dramatically, with a decrease in ellipticity and a decrease in the proportion of α-helices present[26]. Some proteins that are only partially unfolded still have detectable α-helices in the spectra[26]. However, the folded and unfolded profiles of a protein are quite different for different domain structures[27]. As mentioned above, soluble aggregates can be present in a protein sample and can consist of partially unfolded protein as well as folded protein[9]. This may represent the state of AK:DapDc, as there is a degree of aggregation present in purified samples. This would result in a profile like that of Figure 1. Section **5.2**, where some of the α-helices have unfolded giving a less distinct negative band at 222 nm and 208 nm, but the β-strands are still present, with a prominent negative band at 218 nm[27].

Homology modelling using SWISS-MODEL predicted AK:DapDc to have highest homology with the aspartate kinase from *Arabidopsis thaliana* (PDB ID: 2CDQ)[28] and the diaminopimelate decarboxylase from *E. coli* (PDB ID: 1KNW)[29]. There is also a 30 amino

acid sequence between the predicted conserved domains of the two models of AK:DapDc that is potentially unstructured random coil. Both aspartate kinase and diaminopimelate decarboxylase are known to form dimers and tetramers[19, 30]. Aspartate kinase can dimerise either by the ACT regulatory domains interacting, or by the catalytic domains interacting. The *A. thaliana* aspartate kinase dimerises by overlapping the ACT regulatory domains (Figure 1. A). Diaminopimelate decarboxylase dimerises by the TIM-barrel of one monomer interacting with the β-sandwich domain of the other monomer[30], as with Figure 1.C below. There is not tetrameric crystal structure of the diaminopimelate decarboxylase from *E. coli*, but the tetrameric structure for *Mycobacterium tuberculosis* (28% sequence identity with 1KNW) shows that the sides of the TIM-barrels, and part of the β-sandwich domain of each monomer interacts with the corresponding regions of a monomer from the other dimer (Figure 1.D). The dimer interface is between two anti-parallel β-strands from the β-sandwich domains[30].

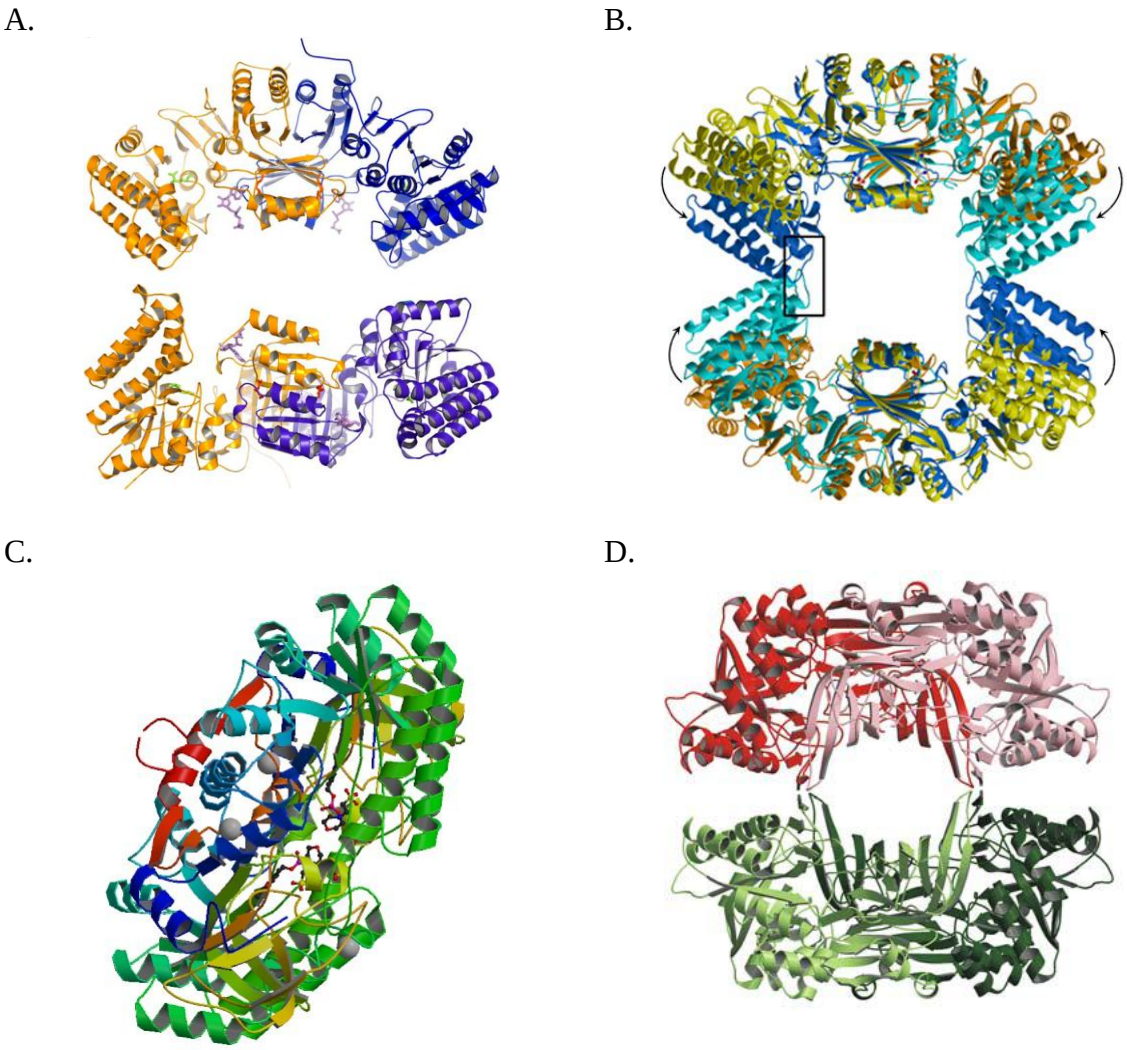


Figure 6.1. Crystal structures showing the dimerisation and tetramerisation of AK and DapDc. A: the dimerisation via the regulatory domains of aspartate kinase from *A. thaliana*, showing the dimer rotated 90°, B: the tetramerisation of aspartate kinase from *E. coli*. The black box shows the region where the two dimers interact, C: The dimerisation of diaminopimelate decarboxylase from *E. coli*, D: the tetramerisation of diaminopimelate decarboxylase from *M. tuberculosis*.

It is unclear how the domains of AK:DapDc would interact and sit in three dimensional space as a monomer or a dimer. The monomer may adopt some of the features of the dimers of both of the template structures predicted by SWISS-MODEL (Figure 2). The ACT domains of the aspartate kinase could overlap the β-sandwich domain of the diaminopimelate decarboxylase function.

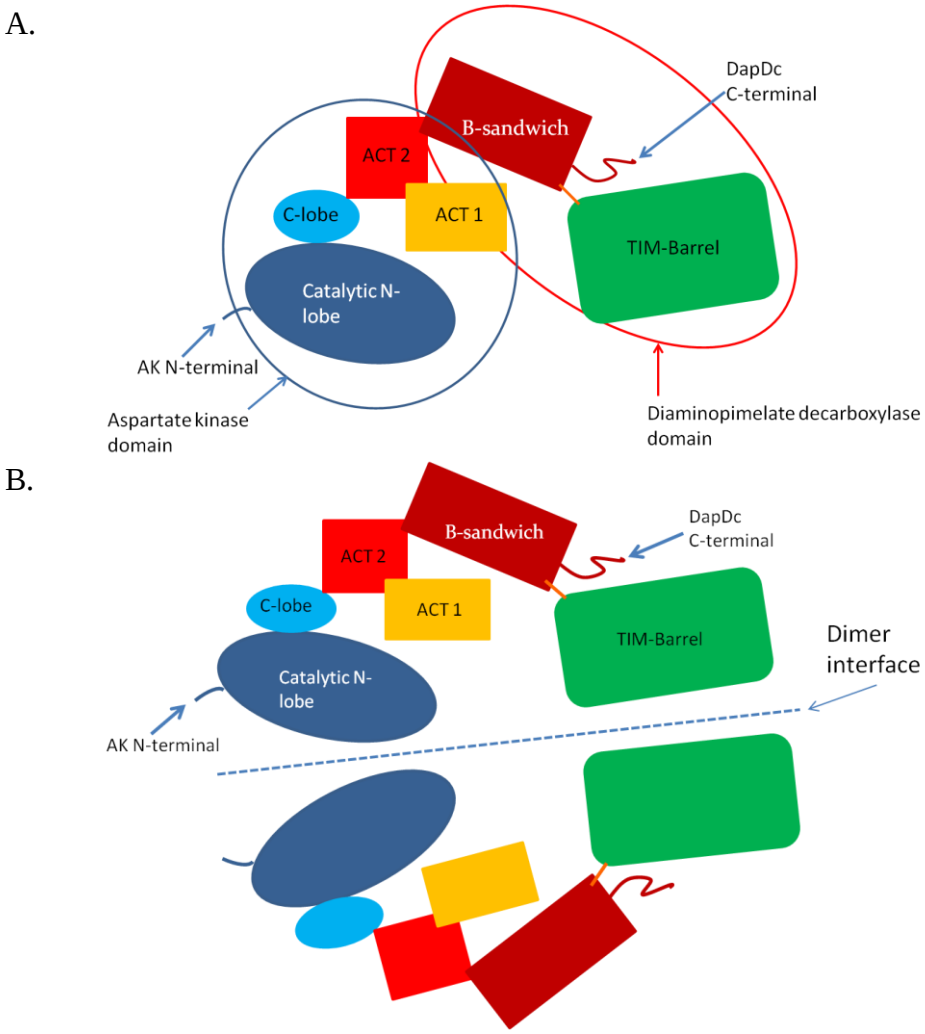


Figure 6.2 Hypothesised AK:DapDc monomeric and dimeric structure based on the dimerisation and tetramerisation of the homology models 1KNW and 2CDQ. A: the hypothesised monomeric structure of AK:DapDc with the ACT 1 and 2 domains overlapping the β-sandwich domain from DapDc. The C-terminal of AK joins up to the N-terminal of DapDc via a linker of unknown conformation. B: The hypothesised dimeric structure of AK:DapDc

The AUC data analysed in this study suggests the presence of both monomer and dimer in solution, but this cannot be confirmed because all samples contained a number of contaminating species and also a small degree of aggregation.

Crystal trays did not result in any crystals. Due to the instability and aggregation present in purified AK:DapDc the protein was beginning to aggregate and precipitate at 10 mg/mL, before being added to the crystallisation solutions. This resulted in most conditions

precipitating the protein (Figure 11, section 5.5). Hopefully if more stable protein could be produced then these same crystal trials would provide more information about crystallising AK:DapDc. The AK:DapDc purified in this project began precipitating at higher concentrations (~10 M), the concentration of protein used in crystallisation trials could be decreased to 5 M to prevent unwanted aggregation.

6.6 Concluding Remarks

In summary, the AK:DapDc protein from *X. fastidiosa* was successfully overexpressed and purified in this study. Its properties were investigated using several biophysical techniques, but these were affected by sample purity and aggregation. The AK:DapDc samples purified in this study could not catalyse the diaminopimelate decarboxylase reaction. However, the expression, purification and buffer optimisation work done in this project provides a good base of knowledge about studying AK:DapDc.

The next steps for this project, in my view, would be to try co-expressing AK:DapDc with chaperones or expressing the two domains separately. The protein could then be purified more easily. I would like to see this protein crystallised and its structure elucidated by x-ray crystallography as I think this protein could help increase our understanding of bifunctional proteins. AK:DapDc is a very interesting protein and further research could aid in the control of *X. fastidiosa*.

References

1. Davis, M. J., Purcell, A. H., and Thomson, S. V. (1978) Pierce's Disease of Grapevines: Isolation of the Causal Bacterium, *Science 199*, 75-77.
2. Baumgartner, K., and Warren, J. G. (2005) Persistence of Xylella fastidiosa in riparian hosts near northern California vineyards, *Plant Disease 89*, 1097-1102.
3. New Zealand Winegrowers. (2013) New Zealand Winegrowers Annual Report 2013.
4. Statistics New Zealand. (2007) Statistics New Zealand - Agricultural Production Census 2007, Statistics New Zealand, Wellington.
5. Lisa Ward, S. H., and Gerard Clover. (2010) Development of a LAMP assay for Xylella fastidiosa, MAF Biosecurity NZ.
6. Costa, H. S., Raetz, E., Pinckard, T. R., Gispert, C., Hernandez-Martinez, R., Dumenyo, C. K., and Cooksey, D. A. (2004) Plant hosts of Xylella fastidiosa in and near southern California vineyards, *Plant Disease 88*, 1255-1261.
7. Dogovski, C., Atkinson, S. C., Dommaraju, S. R., Hor, L., Dobson, R., Hutton, C., Gerrard, J. A., and Perugini, M. (2009) Lysine biosynthesis in bacteria: an unchartered pathway for novel antibiotic design, *Encyclopedia of life support systems 11*, 116-136.
8. Makrides, S. C. (1996) Strategies for achieving high-level expression of genes in Escherichia coli, *Microbiological Reviews 60*, 512-&.
9. Nomine, Y., Ristriani, T., Laurent, C., Lefevre, J. F., Weiss, E., and Trave, G. (2001) A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein, *Protein Engineering 14*, 297-305.
10. Robinson, M., Lilley, R., Little, S., Emtage, J. S., Yarranton, G., Stephens, P., Millican, A., Eaton, M., and Humphreys, G. (1984) Codon usage can affect efficiency of translation of genes in Escherichia-coli, *Nucleic Acids Research 12*, 6663-6671.
11. de Marco, A., Vigh, L., Diamant, S., and Goloubinoff, P. (2005) Native folding of aggregation-prone recombinant proteins in Escherichia coli by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones, *Cell Stress & Chaperones 10*, 329-339.
12. Terpe, K. (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems, *Applied Microbiology and Biotechnology 72*, 211-222.
13. Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures, *Protein Expression and Purification 41*, 207-234.
14. Houry, W. A., Frishman, D., Eckerskorn, C., Lottspeich, F., and Hartl, F. U. (1999) Identification of in vivo substrates of the chaperonin GroEL, *Nature 402*, 147-154.
15. Bignon, C., Li, C. Q., Lichiere, J., Canard, B., and Coutard, B. (2013) Improving the soluble expression of recombinant proteins by randomly shuffling 5 ' and 3 ' coding-sequence ends, *Acta Crystallographica Section D-Biological Crystallography 69*, 2580-2582.
16. Ryzewski, C., and Takahashi, M. T. (1975) Cobalt (III) labeled Aspartokinase-Homoserine Dehydrogenase of Escherichia-coli, *Biochemistry 14*, 4482-4486.

17. Seabrook, S. A., and Newman, J. (2013) High-Throughput Thermal Scanning for Protein Stability: Making a Good Technique More Robust, *Acs Combinatorial Science 15*, 387-392.
18. Xu, H. Y., West, A. H., and Cook, P. F. (2006) Overall kinetic mechanism of saccharopine dehydrogenase from Saccharomyces cerevisiae, *Biochemistry 45*, 12156-12166.
19. Kotaka, M., Ren, J., Lockyer, M., Hawkins, A. R., and Stammers, D. K. (2006) Structures of R- and T-state Escherichia coli aspartokinase III - Mechanisms of the allosteric transition and inhibition by lysine, *Journal of Biological Chemistry 281*, 31544-31552.
20. Chassagnole, C., Rais, B., Quentin, E., Fell, D. A., and Mazat, J. P. (2001) An integrated study of threonine-pathway enzyme kinetics in Escherichia coli, *Biochemical Journal 356*, 415-423.
21. Xu, H., West, A. H., and Cook, P. F. (2007) Determinants of Substrate Specificity for Saccharopine Dehydrogenase from Saccharomyces cerevisiae†, *Biochemistry 46*, 7625-7636.
22. Rosner, A. (1975) Control of Lysine biosynthesis in Bacillus-subtilis - Inhibition of Diaminopimelate decarboxylase by Lysine, *Journal of Bacteriology 121*, 20-28.
23. Bohm, G., Muhr, R., and Jaenicke, R. (1992) QUANTITATIVE-ANALYSIS OF PROTEIN FAR UV CIRCULAR-DICHROISM SPECTRA BY NEURAL NETWORKS, *Protein Engineering 5*, 191-195.
24. Andrade, M. A., Chacon, P., Merelo, J. J., and Moran, F. (1993) EVALUATION OF SECONDARY STRUCTURE OF PROTEINS FROM UV CIRCULAR-DICHROISM SPECTRA USING AN UNSUPERVISED LEARNING NEURAL-NETWORK, *Protein Engineering 6*, 383-390.
25. Provencher, S. W., and Glockner, J. (1981) ESTIMATION OF GLOBULAR PROTEIN SECONDARY STRUCTURE FROM CIRCULAR-DICHROISM, *Biochemistry 20*, 33-37.
26. Greenfield, N. J. (2006) Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions, *Nature Protocols 1*, 2527-2535.
27. Greenfield, N. J. (2006) Using circular dichroism spectra to estimate protein secondary structure, *Nature Protocols 1*, 2876-2890.
28. Mas-Droux, C., Curien, G., Robert-Genthon, M., Laurencin, M., Ferrer, J.-L., and Dumas, R. (2006) A Novel Organization of ACT Domains in Allosteric Enzymes Revealed by the Crystal Structure of Arabidopsis Aspartate Kinase, *The Plant Cell Online 18*, 1681-1692.
29. Levdikov, V., Blagova, L., Bose, N., Momany, C. Diaminopimelate Decarboxylase uses a Versatile Active Site for Stereospecific Decarboxylation, *To be published*.
30. Weyand, S., Kefala, G., Svergun, D. I., and Weiss, M. S. (2009) The three-dimensional structure of diaminopimelate decarboxylase from Mycobacterium tuberculosis reveals a tetrameric enzyme organisation, *Journal of structural and functional genomics 10*, 209-217.

Appendix I: Modified Buffers for DSF Buffer Screening

Table 6.1 - Modified buffers developed in this study for use in DSF buffer screening, adapted from the protocol described in Seabrook & Newman (2013)[1].

Position	Contents	pH	[NaCl]
A1, A2, A3	Lysozyme (+ve control)	8	50mM
A4, A5, A6	Water control	n/a	200mM
A7, A8, A9	Water control	n/a	500mM
A10, A11, A12	Protein control		n/a
B1, B2, B3	Sodium acetate	4.0	200mM
B4, B5, B6	Sodium acetate	4.0	500mM
B7, B8, B9	Sodium acetate	5.0	200mM
B10, B11, B12	Sodium acetate	5.0	500mM
C1, C2, C3	MES	6.0	200mM
C4, C5, C6	MES	6.0	500mM
C7, C8, C9	Citrate	4.0	200mM
C10, C11, C12	Citrate	4.0	500mM
D1, D2, D3	Citrate	5.0	200mM
D4, D5, D6	Citrate	5.0	500mM
D7, D8, D9	Citrate	6.0	200mM
D10, D11, D12	Citrate	6.0	500mM

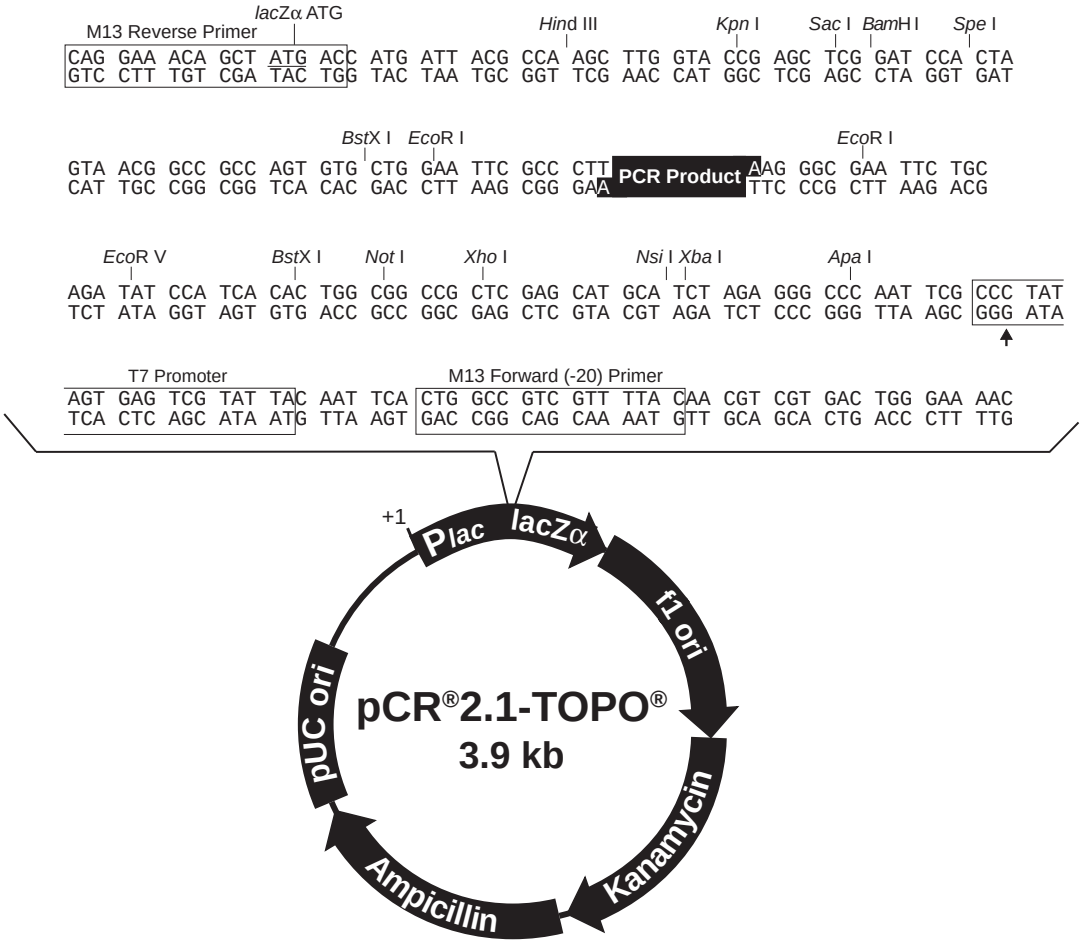
Position	Contents	pH	[NaCl]
E1, E2, E3	Sodium phosphate	6.0	200mM
E4, E5, E6	Sodium phosphate	6.0	500mM
E7, E8, E9	Sodium phosphate	7.0	200mM
E10, E11, E12	Sodium phosphate	7.0	500mM
F1, F2, F3	Bis-Tris Propane	7.0	200mM
F4, F5, F6	Bis-Tris Propane	7.0	500mM
F7, F8, F9	Bis-Tris Propane	8.0	200mM
F10, F11, F12	Bis-Tris Propane	8.0	500mM
G1, G2, G3	Bis-Tris Propane	9.0	200mM
G4, G5, G6	Bis-Tris Propane	9.0	500mM
G7, G8, G9	Tris	7.0	200mM
G10, G11, G12	Tris	7.0	500mM
H1, H2, H3	Tris	8.0	200mM
H4, H5, H6	Tris	8.0	500mM
H7, H8, H9	Sample Benchmark	n/a	n/a
H10, H11, H12	Dye control	n/a	n/a

References

1. Seabrook, S. A., and Newman, J. (2013) High-Throughput Thermal Scanning for Protein Stability: Making a Good Technique More Robust, Acs Combinatorial Science 15, 387-392.

Appendix II: Plasmid Maps

pCR2.1-TOPO, pBluescript-SK^{+}, pET30ΔSE plasmid maps follow. pET30GST is a modified version of pET30ΔSE with the N-terminal His-tag replaced by a GST tag.

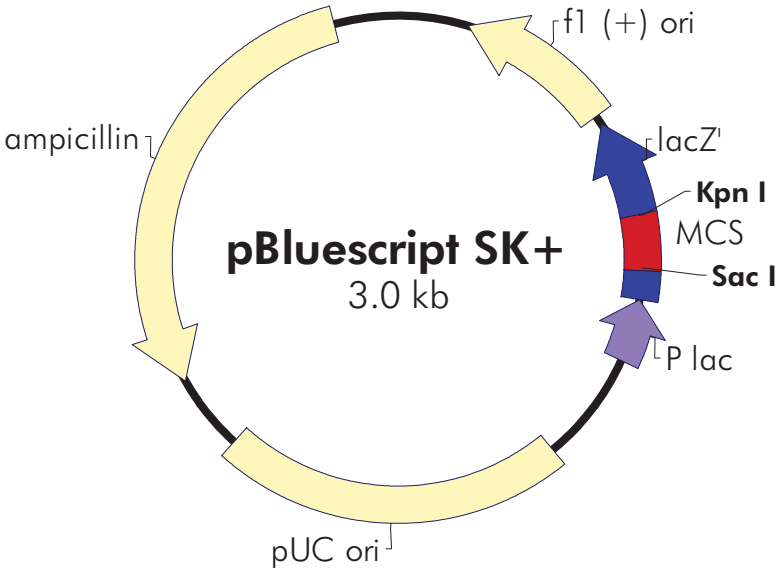


Comments for pCR®2.1-TOPO®
3931 nucleotides

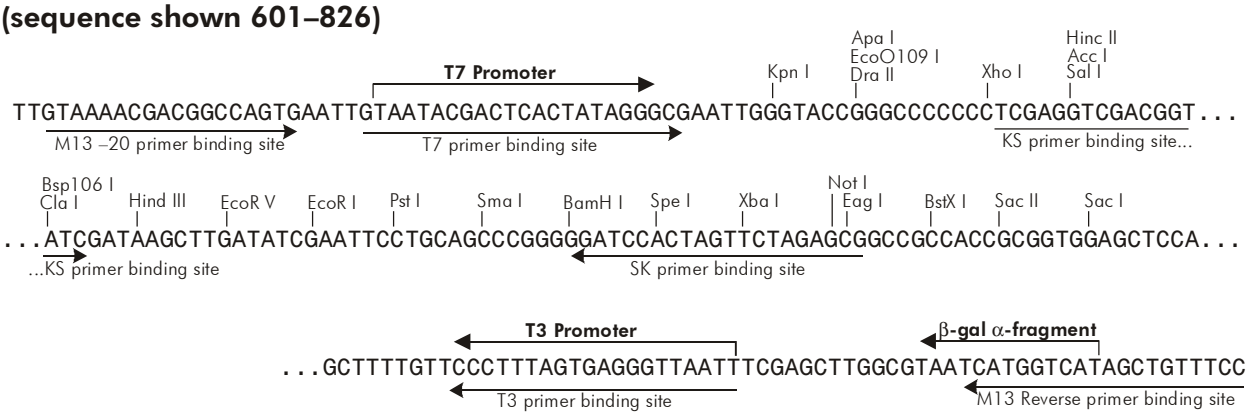
*Lac*Zα fragment: bases 1-547
M13 reverse priming site: bases 205-221
Multiple cloning site: bases 234-357
T7 promoter/priming site: bases 364-383
M13 Forward (-20) priming site: bases 391-406
f1 origin: bases 548-985
Kanamycin resistance ORF: bases 1319-2113
Ampicillin resistance ORF: bases 2131-2991
pUC origin: bases 3136-3809

Invitrogen™
life technologies

f1 (+) origin 138–444
β-galactosidase α-fragment 463–816
multiple cloning site 653–760
***lac* promoter 817–938**
pUC origin 1158–1825
ampicillin resistance (*bla*) ORF 1976–2833



pBluescript SK (+/–) Multiple Cloning Site Region (sequence shown 601–826)



pET-30a-c(+) ΔStag-Enetero

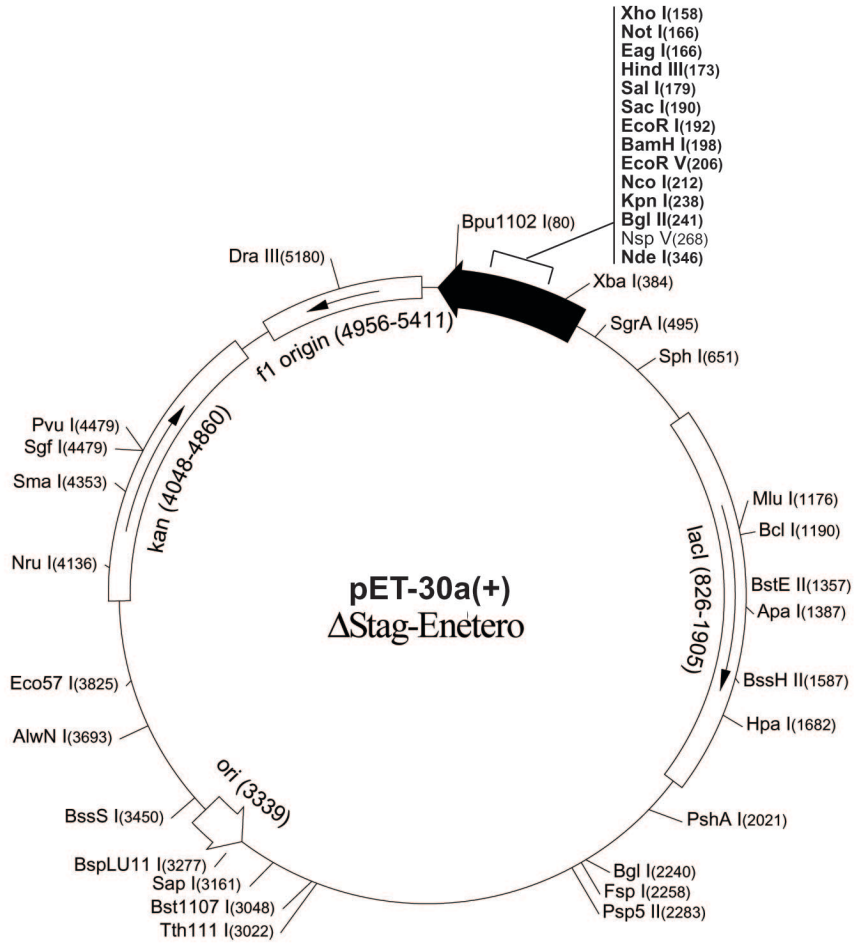
	Cat. No.
pET-30a DNA	69909-3
pET-30b DNA	69910-3
pET-30c DNA	69911-3

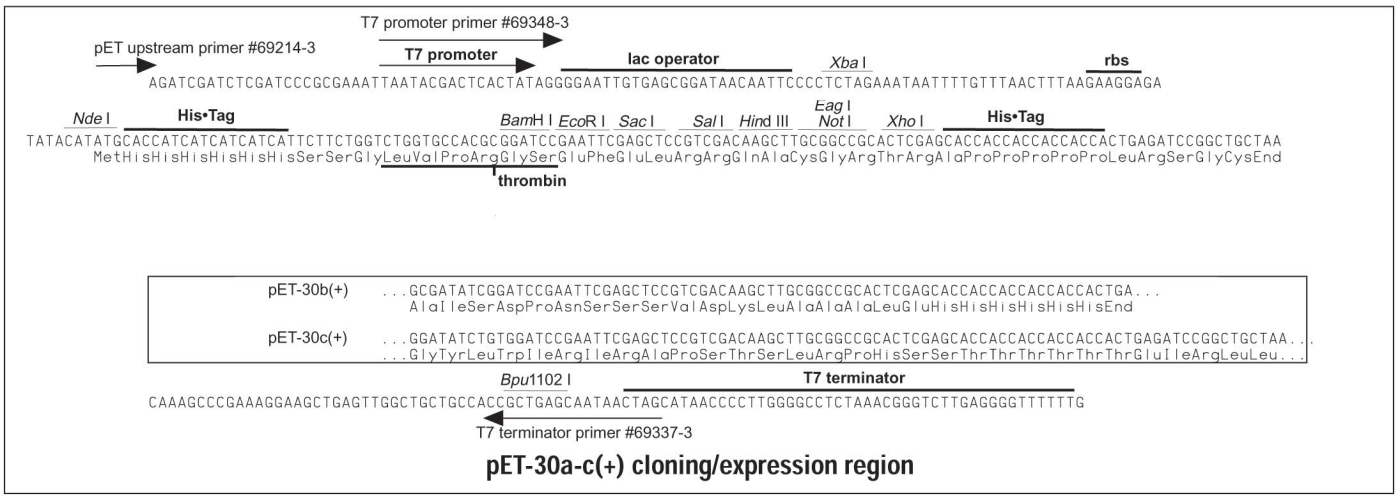
The pET-30a-c(+) vectors carry an N-terminal His•Tag®/thrombin/S•Tag™/enterokinase configuration plus an optional C-terminal His•Tag sequence. Unique sites are shown on the circle map. Note that the sequence is numbered by the pBR322 convention, so the T7 expression region is reversed on the circular map. The cloning/expression region of the coding strand transcribed by T7 RNA polymerase is shown below. The f1 origin is oriented so that infection with helper phage will produce virions containing single-stranded DNA that corresponds to the coding strand. Therefore, single-stranded sequencing should be performed using the T7 terminator primer (Cat. No. 69337-3).

pET-30a(+) sequence landmarks

T7 promoter	419-435
T7 transcription start	418
His•Tag coding sequence	327-344
S•Tag coding sequence	249-293
Multiple cloning sites	
(*Nco* I - *Xho* I)	158-217
His•Tag coding sequence	140-157
T7 terminator	26-72
lacI coding sequence	826-1905
pBR322 origin	3339
Kan coding sequence	4048-4860
f1 origin	4956-5411

The maps for pET-30b(+) and pET-30c(+) are the same as pET-30a(+) (shown) with the following exceptions: pET-30b(+) is a 5421bp plasmid; subtract 1bp from each site beyond *Bam*H I at 198. pET-30c(+) is a 5423bp plasmid; add 1bp to each site beyond *Bam*H I at 198.





pET-30a-c(+) cloning/expression region